Studying neural selectivity for motion using high-field fMRI

Alex Beckett, MSc.

Thesis submitted to The University of Nottingham
for the degree of Doctor of Philosophy

February 2013

Abstract

Functional magnetic resonance imaging (fMRI) offers a number of opportunities to non-invasively study the properties of the human visual system. Advances in scanner technology, particularly the development of high-field scanners, allow improvements in fMRI such as higher resolution and higher signal to noise ratio (SNR). We aimed to examine what these advances in scanner technology, combined with novel analysis techniques, can tell us about the processing of motion stimuli in the human visual cortex.

In Chapter 3 we investigated whether high-resolution fMRI allows us to directly study motion-selective responses in MT+. We used event-related and adaptation methods to examine selectivity for coherent motion and selectivity for direction of motion, and examined the potential limitations of these techniques.

One particular analysis technique that has been developed in recent years uses multivariate methods to classify patterns of activity from visual cortex. In Chapter 4 we investigated these methods for classifying direction of motion, particularly whether successful classification responses are based on fine-scale information such as the arrangement of direction-selective columns, or a global signal at a coarser scale.

In Chapter 5 we investigated multivariate classification of non-translational motion (e.g. rotation) to see how this compared to the classification of translational motion. The processing of such stimuli have been suggested to be free from the large-scale signals that may be involved in other stimuli, and therefore a more powerful tool for studying the neural architecture of visual cortex.

Chapter 6 investigated the processing of plaid motion stimuli, specifically 'pattern' motion selectivity in MT+ as opposed to 'component' motion selectivity. These experiments highlight the usefulness of multivariate methods even if the scale of the signal is unknown.

Parts of the work discussed in Chapter 4 were published in the following article:

Beckett A, Peirce J, Sanchez-Panchuelo R, Francis S, & Schluppeck D 'Contribution of large scale biases in decoding of direction-of-motion from high-resolution fMRI data in human early visual cortex.' *NeuroImage* 2012

Acknowledgements

Firstly, I would like to thank my supervisor Denis Schluppeck for first giving me the opportunity to work with fMRI, and for the extensive help and feedback he gave during the writing of this thesis.

I would also like to thank Jon Peirce for his help and feedback for the work that went into this thesis.

In addition I would like to thank Julien Beslie and Rosa Sanchez for their help with the methodology and data collection in these experiments. I would also like to thank the Nottingham Visual Neuroscience Group for providing such a great working environment.

Finally, I would also like to offer a huge thank you to Molly Simmonite. I couldn't have done this without her.

Contents

Chapter 1

Introduction

The human visual system takes in a huge amount of raw information, in the form of patterns of light hitting the back of the eye, and extracts from this useful information about our surroundings. Understanding the process by which this analysis is performed has been a key aim of psychology and neuroscience since their inception. Among the tools available to psychologists and neuroscientists are behavioural experiments with psychophysics and visual illusions, examining the neuropsychological effects of brain damage on vision, and measuring the activity of single cells in animals in response to visual stimuli.

The recent development of non-invasive techniques for measuring and analysing neural information has been key in understanding how the visual system ultimately transforms the patterns of light entering the eye into a neural signal that allows us to understand and interact with the world around us. The most consistently used neuroimaging technique for human volunteers has been Functional Magnetic Resonance Imaging (fMRI), with a huge increase in the number of published papers using the technique since its development in the early 1990's (Ogawa et al., 1990). In this thesis I will examine what recent technical developments in this field can tell us about the analysis of visual motion information in the human visual cortex.

FMRI has allowed studies of the processing performed in each stage of the human visual system to be related to evidence from electro-physiology in animals, neuropsychological studies of human subjects after brain damage, and behavioural studies in healthy subjects. The following sections summarise the current literature on the function of the visual system, and how the evidence from different research methods complement and differ from one another.

1.1 The Visual System

The processing of visual information begins with light entering the eye via the pupil and hitting the retina on the back of the eye, and continues on through the visual cortex and beyond. In the following sections I will briefly summarise the pathway from the retina to the visual cortex, and examine the different aspects of analysis at each stage.

1.1.1 From Retina to Brain

The initial stage of processing for visual information occurs when light falls on the retina, the array of photoreceptors on the back of eye, having been focussed (and inverted) by the lens of the eye. In vertebrates, the output of these photoreceptors is collected and combined by retinal ganglion cells, with a large amount of processing being done at this initial stage (Callaway, 2005; Lettvin et al., 1959). Whilst the photoreceptors in the retina simply change their level of response depending on the light that falls on them, the retinal ganglion cells are more selective in their responses. It is worth noting that a significant amount of processing is done at the very first level of processing in the retina, and important information about a visual scene is extracted at this very early stage, well before the visual cortex itself (although the details of this are beyond the scope of this thesis). The area of retina (and hence visual field) which will cause a retinal ganglion cell to fire if stimulated is called the 'receptive field' (RF) for that cell, and many retinal ganglion cells in vertebrates, for examples those in the cat retina, have a 'centre-surround' RF arrangement, with a central region that either excites or inhibits the cell in response to light, and surrounding ring with the opposite sensitivity (Kuffler, 1953). Light falling across the whole RF will cause the cell to fire very weakly, whereas light falling only on the excitatory centre (if the cell is 'on-centre') will cause the cell to fire rapidly (if the cell is 'off-centre', light falling on the surround only will cause rapid cell firing). Therefore these cells are sensitive to contrast, discontinuities in the distribution of light corresponding to edges, rather than simply to different levels of illumination.

The axons of the majority of retinal ganglion cells project to the brain along the optic nerves. The optic nerves from each eye meet at the optic chiasm (Figure 1.1), where the information is combined and split depending on its origin on the retina. Fibres originating from the nasal part of the retina cross over to the other side of the brain, whilst fibres originating at the temporal side of the retina continue on the same side. The result of this *decussation* is to split the visual field into a left and right portion, with the left visual field (from both eyes) being processed by the right side of the brain and

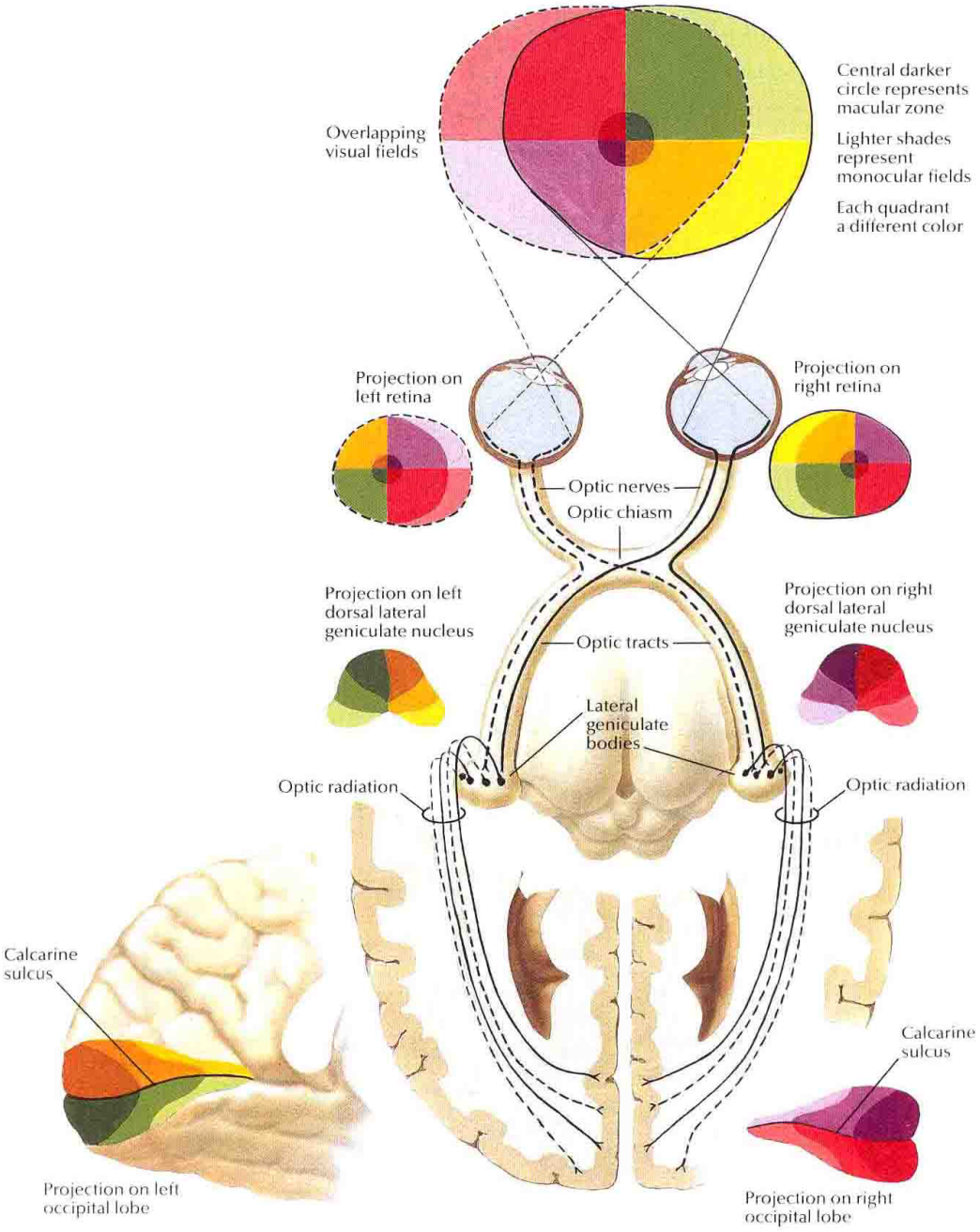

Figure 1.1: A diagram indicating the path taken by visual information from the eye to the visual cortex, indicating how the visual field is represented at each stage in each cortical hemisphere. Adapted from Netter (2010)

the right visual field being processed by the left side of the brain (Figure 1.1).

Following the optic chiasm, the optic nerve is referred to as the optic tract. The majority of the fibres which make up the optic tract terminate at the Lateral Geniculate Nucleus (LGN) in the thalamus (Figure 1.1); the rest terminate in the midbrain, primarily at the Superior Colliculus, as well as the suprachiasmatic nucleus.

In many primates, including humans, the LGN is subdivided into a number of layers, with different layers receiving input from different populations of retinal cells, with preferences for different kinds of stimuli (Wiesel and Hubel, 1966). In addition, each layer receives input from only one eye. The uppermost layers, known as parvocelluar layers, receive input from a class of cells known as midget retinal ganglion cells whilst the bottom layers, the magnocellular layers, receive input from parasol cells. These retinal ganglion cells have different receptive field properties, and these two kinds of layers form the beginning of two segregated visual pathways that continue through the visual system (Livingstone and Hubel, 1988). The two visual streams process different kinds of stimuli, with the parvocellular (P) pathway favouring high spatial frequency and colour information, and the magnocellular (M) pathway carrying coarser spatial frequency and motion information. Cells in between the magno- and parvocellular layers receive input from bistratified retinal ganglion cells and represent a third processing stream, the koniocellular stream, whose perceptual specialization is unclear at this point. Each layer of the LGN contains a full representation of the contralateral visual field (Figure 1.1), and axons projecting from neighbouring parts of the retina terminate at neighbouring geniculate cells, creating an ordered map of the retina. This representation that preserves the topography of the retina is known as a retinotopic map, and is a feature of many mammalian brains. The layers of the LGN are arranged in such a way that the retinotopic maps of each layer are aligned. The receptive fields of neurons in mammalian LGN closely resemble those of retinal ganglion cells in terms of on-off surround (Hubel and Wiesel, 1961). Although the exact function of the LGN is still debated (Callaway, 2005), its separation of signals from the retina in terms of function and origin is believed to set-up a similar segregation in visual cortex. The LGN projects to both the visual cortex and additionally to the superior colliculus, a nucleus involved with the control of eye movements.

The LGN projects to the cortex via the optic radiation, which terminates in layer 4 of the primary visual cortex (V1). The retinotopic representation established in the LGN is maintained in V1: different areas of the contralateral visual field are mapped in an orderly fashion in area V1 of each cortical hemisphere, with adjacent points in the visual field being processed by adjacent neurons in cortex (Figure 1.1). One key feature

of the retinotopic map that is present at the LGN and emphasised in visual cortex is that the fovea tends to be overrepresented compared to the periphery of the retina, with larger receptive fields in the periphery and smaller at the fovea. This property is known as cortical magnification. Retinotopic organization persists in many visual areas beyond V1, and this feature is exploited in identifying and defining visual areas with fMRI in individual subjects. A full discussion of the methodology for retinotopic mapping is given in section 2.2.

1.1.2 V1

V1 lies in the calcarine sulcus at the posterior pole of the occipital cortex (Figure 1.1). V1 has a well defined representation of the contralateral visual field, organised retinotopically, with adjacent points of the visual field represented at adjacent locations on the cortical surface. As with ganglion cells in the retina, cortical cells do not simply respond to levels of contrast, but are selective for certain properties of an image. The properties of V1 cells were first extensively studied in animals by Hubel and Wiesel (1959, 1963, 1969). One key finding was that a population of cells in cat primary visual cortex, which they named 'simple cells', would show preferential activity for bars of light oriented at a specific angle. This selectivity is due to the shape of the cells' receptive fields, which have elongated On and Off regions with a given orientation (Hubel and Wiesel, 1959). These cells will only fire when the dark and light portions of an oriented bar fall exactly on the correct regions, making them highly selective for position and orientation. A second class of cells, called 'complex cells' by Hubel and Wiesel (1962) has the same selectivity for orientation, but their receptive fields do not have as defined On/Off regions as simple cells, so they respond to a properly oriented stimulus falling anywhere in its receptive field. A third class of cells, known as hypercomplex or 'end-stopped' cells, are sensitive to the length of a stimulus as well as its orientation, and will reduce their response if the stimulus exceeds the length of the receptive field (Hubel and Wiesel, 1965). These sub-classes of visual cells have also been demonstrated in non-human primates (Hubel and Wiesel, 1968).

A further finding by Hubel and Wiesel (1959) was that cells with the same orientation selectivity were grouped together perpendicular to the cortical surface, leading to the development of the idea of 'orientation columns' in primary visual cortex. Cortical columns were initially identified in somatosensory cortex of the cat (Mountcastle et al., 1957), where cells perpendicular to the cortical surface had sensitivity to the same kind of tactile stimulation. Hubel and Wiesel (1959) found that the preferred orientation of cells in cat visual cortex was constant as the recording electrode was pushed perpen-

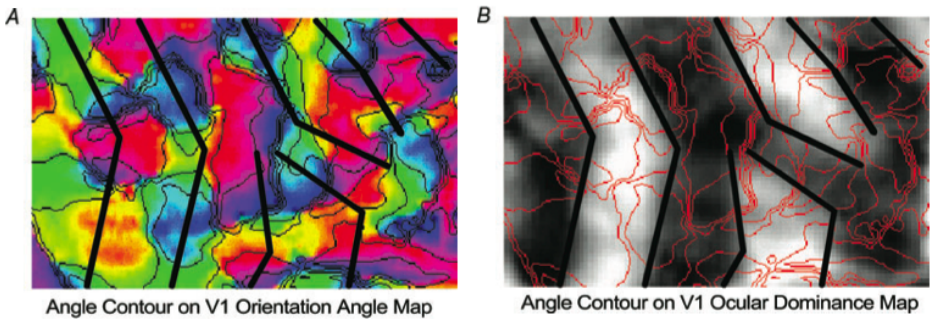

Figure 1.2: Orientation Columns (A) and ODCs (B) measured from primate V1 using optical imaging. Colour coding in A indicates preferred orientation (horizontal = blue, 45° = red, vertical = yellow, 135° = green). Colour coding in B indicates preference for stimulated eye (dark = left eye, light = right eye). Dark lines in both figures indicate borders between ODCs, thin-lines indicate iso-orientation contours. The two dimensions can be seen to run broadly orthogonal to each other. Taken from Ts'o et al. (2009)

dicularly through the cortical surface, but varied regularly as the electrode progressed obliquely. Similar results were found in macaque visual cortex (Hubel and Wiesel, 1974), leading to the formulation of a model of orientation 'slabs' arranged across the cortical surface, with adjacent slabs having slightly shifted orientation preferences. Hubel and Wiesel (1974) coined the term 'hypercolumn' to describe an area of cortex containing a set of columns with the full range of orientation preferences. As well as preferences for stimulus orientation, Hubel and Wiesel (1962, 1969) also showed that cells in visual cortex of cats and macaques have a preference for stimulation through one eye versus the other, and that cells with similar eye preference are also arranged on the cortical surface into 'Ocular Dominance Columns' (ODCs). These two observations led to the development of what came to be known as the 'ice-cube' model, with hypercolumns for orientation and ODCs orthogonal to each other on the cortical surface, with any given area of cortex containing multiple, overlapping columns (Hubel and Wiesel, 1977), hence containing cells tuned across a complete range of values for each domain. This block of tissue was referred to as a 'module', and set forth the idea that these discrete units were responsible for analysing the visual field fully across these domains at a given retinotopic location.

Further evidence of an ordered arrangement of selective neurons came from optical imaging, which uses the light reflected from an exposed cortical surface to measure neural activity. This allows the preferences of a large number of cells in the visual cor-

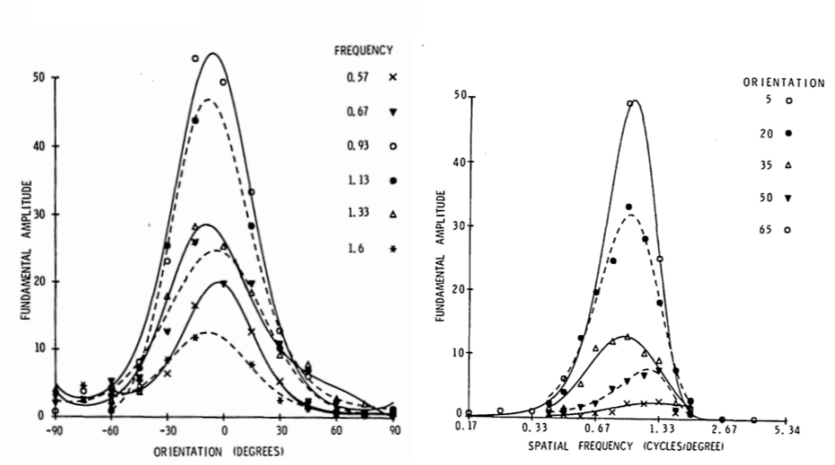

Figure 1.3: Example of a simple cell in cat visual cortex with orientation (left) and spatial frequency (right) tuning, demonstrating the bell shaped tuning curve in both instances. Adapted from Webster and De Valois (1985)

tex to be measured simultaneously, and allows the direct visualization of the layout of neuronal preference maps on the cortical surface (Figure 1.2). Whilst several studies indicated that orderly arrangement of cells with similar preferences for orientation and stimulated eye existed in cat and primate visual cortex and were broadly orthogonal (Grinvald et al., 1986; Ts'o et al., 2009), some revision of the Hubel and Wiesel (1977) model was necessary. For example the addition of 'pin-wheel' arrangement of preferred orientation (Bonhoeffer and Grinvald, 1991), where the preferred orientation of the cells progressed radially around a centre point instead of along the cortical surface (although some have argued that this feature of orientation maps from optical imaging is simply an artefact caused by draining veins). The concept of a 'hypercolumn' however has been more elusive, with a periodic repetition of a 'module' containing neural mechanisms for a full analysis of visual space often difficult to establish (Bartfeld and Grinvald, 1992).

In humans, the existence of columns was initially demonstrated using post-mortem cytochrome oxidase (CO) tissue staining (which stains cells based on their metabolic activity) in the visual cortex of patients with monocular vision loss, leaving ODCs for the missing eye lighter than those for the healthy eye (Adams et al., 2007). Recently, the existence of ODCs and orientation columns in human V1 has been demonstrated non-invasively using high-resolution fMRI at 7T (Yacoub et al., 2007, 2008).

Cells in V1 also display preferences for additional stimulus properties, for example spatial frequency (relating to the level of detail in an image) (Campbell et al., 1969;

De Valois et al., 1982a). Cortical neurons demonstrate selectivity for both a given orientation and a given spatial frequency, generally showing a bell shaped tuning curve to both properties (Figure 1.3). Neurons such as this can be thought of a acting as a set of spatial frequency filters at different orientations, essentially performing a crude form of Fourier analysis of an image. Whilst columnar architecture for orientation has been demonstrated in both animals (Hubel and Wiesel, 1962, 1974) and humans (Yacoub et al., 2008), an ordered representation for spatial frequency has not been.

A subset of cells in cat and monkey V1 have been shown to be selective for direction of motion as well as orientation (De Valois et al., 1982b; Hubel and Wiesel, 1962), in that neurons increase their activity for a contour of their preferred orientation moving in a given direction of motion. An ordered map of direction preference, as for orientation, has been demonstrated in early visual cortex for some animals (Welicky et al., 1996), although not in primate V1 (Lu et al., 2010) where only *axis of direction* columns could be demonstrated. Cells in V1 with motion selectivity primarily project to areas thought to be involved in motion processing, such as area MT/V5 , both directly and via areas such as V2 and V3 (Maunsell and van Essen, 1983).

1.1.3 Visual Cortical Pathways

V1 projects directly to a number of other visual areas, as well as indirectly to a number of others via V2. One key feature of these cortical projections is a segregation into two visual pathways. The majority of connections from V1 (via V2 and V4) project ventrally towards the temporal lobe, and are primarily made up of projections from the P pathway in the LGN. The remainder of connections project dorsally towards the parietal lobe, and are primarily of the M pathway. This continues the two visual streams established at the LGN, and suggests that these two streams have distinct functions based on the specialization for form and motion in the two pathways (Livingstone and Hubel, 1988). Although the segregation between magnocellular and parvocellular projections in the two streams may not be absolute (Maunsell et al., 1990), the idea of two cortical streams specialized for different aspects of visual processing has also been supported by electrophysiology and lesion studies.

Ungerleider and Mishkin (1982) named these two projections the *dorsal* and *ventral* streams based on the direction of their projections from V1, and from work with macaque lesions hypothesised the dorsal stream as the 'where' pathway, concerned with spatial awareness, and the ventral stream as the 'what' pathway, dealing with the recognition of objects. Evidence for such an interpretation comes from a 'double dissociation' in human studies after brain damage, with lesions of posterior parietal cortex leading

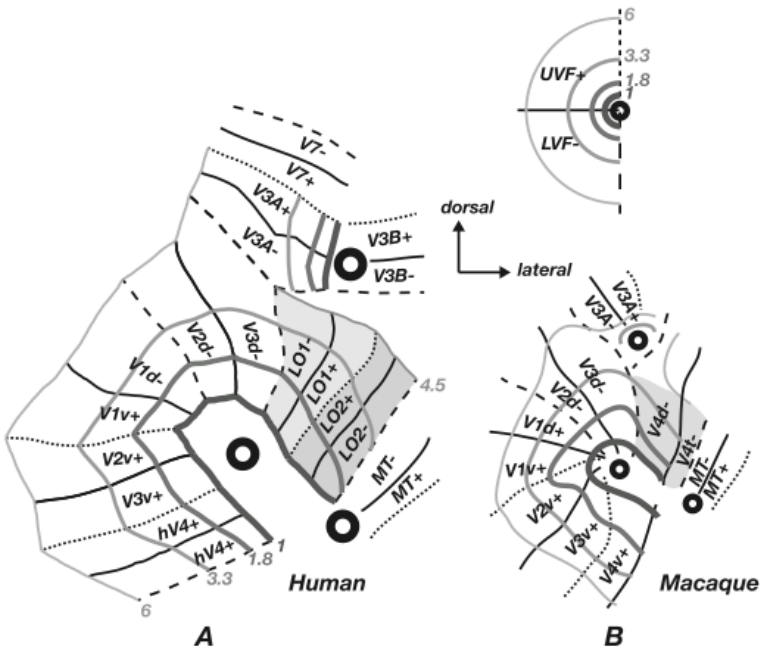

Figure 1.4: Representations of the visual field (Top Right) in the visual cortex of human (A) and macaque (B). Representations of cortex not to scale. Ventral V1-3 contain representations of the contralateral Upper Visual Field (UVF), dorsal V1-3 contain representations of the contralateral lower visual field (LVF). Some areas beyond contain a representation of the full contralateral visual field. Figure taken from Larsson and Heeger (2006).

to optic ataxia (a disorder involving failures of hand-eye coordination), and lesions of ventral visual areas leading to visual agnosia (an impairment of recognition of visually presented objects) (see Milner and Goodale (2008) for a review). Goodale and Milner (1992) presented the split in terms of 'vision for action' in the dorsal stream, and 'vision for perception' in the ventral stream. The independence and separation of the dorsal and ventral streams has been questioned in recent years (Schenk and McIntosh, 2010), and the picture emerging seems to be of a relative rather than an absolute specialization for different aspects of vision, with a large amount of interaction between the areas. However, the two streams hypothesis provides a useful framework to consider the different types of processing done in each visual area.

1.1.4 Extrastriate Cortical Areas

The extrastriate areas that V1 projects to (both indirectly and directly) are defined by having an ordered retinoptopic mapping, preserving the orderly representation that exists in LGN and V1 (Figure 1.4). Receptive fields in these area also tend to be larger

than those in V1. In addition, some extrastriate areas have been shown to demonstrate a preference for specific kinds of visual processing. The next section reviews some of the evidence of the different functional properties of these areas.

V2

Area V2 in each hemisphere comprises two areas, dorsal and ventral of V1 respectively, each with a map of a quadrant of the contralateral visual field (Figure 1.4). V2 receives strong connections from V1, as well as sending many feedback connections to this area. V2 also projects to areas V3, V4, and V5/MT. Cells in V2 show many tuning properties similar to V1 cells such as orientation, spatial frequency and binocular disparity (Levitt et al., 1994). In addition, V2 cells also have additional properties such as tuning for relative disparity (Thomas et al., 2002) and tuning for more complex interactions of orientations (Hegdé and Van Essen, 2000). This suggests that V2 is responsible for building upon the simple visual processing undertaken in V1 to allow more complex processing. Lesions of this area in the macaque affect performance in complex spatial tasks with no effect on acuity or contrast sensitivity (Merigan et al., 1993).

It has been demonstrated that tuning for disparity follows a columnar arrangement in macaque V2, orthogonal to one for orientation (Ts'o et al., 2009). V2 has a striped organization, with different stripes known as thick, thin and pale depending on their appearance after staining with cytochrome oxidase (Livingstone and Hubel, 1984), and it has been suggested that the different stripes contain cells with functionally distinct properties; for disparity and orientation in the thick stripes, colour in the thin stripe and orientation in the pale stripes (Roe and Ts'o, 1995). This suggests the mechanisms for a full analysis of a point in visual space are more distributed and segregated over a patch of cortex in V2 (Ts'o et al., 2009), compared to V1, where hypercolumns for OD and orientation are expected to overlap and interact. Additionally to the organization for orientation and disparity, a map for preferred direction running orthogonally to preferred orientation has been demonstrated in the thick stripes of macaque V2 (Lu et al., 2010), which are known to project to direction selective areas such as MT and V3A.

V3

An area known as V3 lies anterior of V2 on both dorsally and ventrally (Figure 1.4), which receives input from V1 and V2 and projects to V4 and MT. The exact makeup of this area, including how many areas it is subdivided into and their functional prop-

erties, is still a matter of debate (see Lyon and Connolly (2012) for a recent review). V3 neurons in macaque show tuning for orientation (although broader than the tuning seen in V2), a large proportion show direction selectivity (with some evidence of pattern selectivity) and some evidence for colour selectivity (Gegenfurtner et al., 1997). The dorsal part of V3 contains a representation of the lower visual field only, and the corresponding ventral area containing the upper field is sometimes considered as a separate area known as Ventral Posterior (VP) (Burkhalter and Van Essen, 1986), with selectivity for motion and colour split between the two areas. The extent to which this ventral area is functionally distinct from V3 is still controversial (Lyon and Connolly, 2012; Zeki, 2003). For the purposes of this thesis, the areas bordering V2 that together contain a complete representation of the contralateral visual field were treated as a single visual area.

In macaques, an additional dorsal visual area beyond V3 and separate from V4 was identified on the basis of its retinotopic map, and was named the V3 Accessory area (subsequently known as V3A) (Van Essen and Zeki, 1978). V3A was subsequently shown to be an entirely separate area to V3, although the V3A name was kept (Tootell et al., 1997). One key feature of macaque V3A is that it contains a full representation of both the upper and lower parts of the contralateral visual field, unlike other nearby areas (Van Essen and Zeki, 1978) (Figure 1.4). Macaque V3 was shown to be highly direction and motion selective, whereas macaque V3A was shown to be much less so. An area beyond human dorsal V3 was also identified on the basis of its retinotopy, and was also referred to as V3A owing to the similarity of its location to macaque V3 (Tootell et al., 1997). However, in humans the function of the two areas appear to be reversed, in that V3A shows greater motion and direction selectivity than V3 in fMRI experiments (Tootell et al., 1997). An additional area lateral to human V3A was also identified, containing a full representation of the contralateral visual field and sharing a foveal confluence with V3A separate from that of V1-V3, known as V3B (Smith et al., 1998) (Figure 1.4). The exact function of human V3B is unclear at this point, although it appears to be involved in processing shape information (Zeki, 2003), and it is frequently considered as a single Region of Interest (ROI) combined with V3A. Beyond V3A/B, an addition representation of the contralateral hemifield exists in humans in an area known as V7 (Tootell et al., 1998) (Figure 1.4), although the function of this area is unclear.

V4

In the macaque, V4 lies anterior to V3, and contains a full representation of the contralateral visual field, with the upper and lower quadrants in the ventral and dorsal sections respectively (Figure 1.4). V4 receives input from V2 and V1, as well as regions of temporal cortex and areas within the dorsal stream. V4 was initially suggested as a specialized area for colour processing (Zeki, 1983), based on colour (rather than wavelength) selective receptive fields, with some evidence of a columnar organisation for colour. This specialization for colour was contrasted against the apparent specialization for motion in area MT, suggesting the modularity of processing for different aspects of vision. Studies also demonstrated selectivity for orientation in a majority of V4 neurons, with some evidence that cells in V4 could only be optimally driven using more complex stimuli than the sinusoidal gratings used in studying earlier areas such as V1 (Desimone and Schein, 1987). Lesions of macaque V4 showed slight deficits in certain detection and discrimination tasks, as well as more severe deficits in tasks involving the discrimination of form (Merigan, 1996), suggesting a broader role for this area than simply being the 'colour' area of cortex.

The human homologue of macaque V4 has been difficult to ascertain. One key area of debate has been whether the visual hemifield representation is split dorsally and ventrally, or whether a complete hemifield representation exists within ventral V4, and no dorsal V4 exists in humans (Tootell and Hadjikhani, 2001). Lesions of ventral cortex in humans can lead to achromatopsia (deficit in colour vision) (Zeki, 1990), and perceptual deficits similar to those seen in monkeys, suggesting that the human homologue of macaque area V4 lies in this part of cortex.

V5/MT and MST

It has been suggested that some areas in extrastriate cortex, specifically dorsal/parietal areas, are specialized for motion processing. Direction-selective cells in monkey V1 project to an area of the superior temporal sulcus (STS) known as the middle temporal area (MT), which is itself heavily direction selective (Albright et al., 1984; Dubner and Zeki, 1971; Zeki, 1980). This area, known as either MT or V5 was suggested as an example of a specialized area for motion processing (Zeki, 1974), due to its strong direction selectivity and apparent nonselectivity for other visual properties such as colour. Additionally, lesions of V5/MT of the macaque elevate direction-detection thresholds without affecting contrast detection thresholds (Newsome and Pare, 1988).

Area V5 receives inputs from V1, V2, dorsal V3, LGN and the pulvinar, with the vast

majority of its input coming from layer 4B of V1, which is know to have a higher proportion of direction-selective cells compared to other layers (Maunsell and van Essen, 1983). Direction-selective responses in V5/MT persist after inactivation of V1, indicating a role for sub-cortical inputs in direction selectivity, with residual direction selectivity removed by inactivation of the SC (Gross, 1991). MT has also been suggested to be a key part of the dorsal cortical processing stream (Ungerleider and Mishkin, 1982), and projects to a number of areas in posterior parietal cortex. It also projects to areas known to be sensitive to optic flow (MST) and for controlling eye movements (Frontal Eye Fields). Similar to earlier visual areas, V5 is also retinotopically organized, containing a full representation of the contralateral hemifield (Van Essen et al., 1981) (Figure 1.4 B).

Columns for preferred direction have been demonstrated in macaque MT (Albright et al., 1984; Dubner and Zeki, 1971), with smooth changes in preferred direction across successively sampled cells by the penetrating electrode, punctuated by sudden jumps in preference of 180 degrees. These findings suggested that columns of a given preference sit side by side with columns with the opposite preference, with these co-arranged units responding to opposite directions of motion referred to as axis of motion columns (Albright et al., 1984; Diogo et al., 2003). In addition, a columnar organisation for binocular disparity has been demonstrated in macaque MT (DeAngelis and Newsome, 1999), coexisting with the columnar arrangement for direction. Clustering for speed tuning has also been shown, though not with a columnar arrangement. Combined electrophysiological and behavioural work in primates indicates that activity in V5 is linked to the *perception* of motion: activity in this area predicts perceived direction and stimulation biases perceived direction toward the stimulated direction (Salzman et al., 1992).

Although the proportion of direction selective cells in MT/V5 is higher than earlier visual areas, their tuning for direction, speed and disparity does not differ greatly from cells in V1 that project to this area (Movshon and Newsome, 1996). One way in which MT cells are differentiated from direction selective cells in V1 is by their larger receptive fields, leading to suggestions that MT neurons may allow motion be detected over larger displacements, mirroring the difference between 'long-range' and 'short-range' motion processes suggested by psychophysics (Mikami et al., 1986). However, the upper limits of spatial displacement that MT cells are sensitive to are similar to those in V1, despite the much larger RFs in MT (Churchland et al., 2005). An additional suggestion for the functional role of MT is that it is instrumental in computing the motion of whole objects, contrasted with a focus on local motion computation in earlier areas such as V1. A large proportion of cells in MT/V5 are selective for pattern direction,

rather than component direction (Movshon et al., 1985). For example, for a drifting plaid stimulus formed from two drifting gratings, component selective cells will respond to whichever grating more closely matches their preferred orientation, whilst a pattern selective cell will respond to the motion of the plaid itself, regardless of the underlying components.

A number of additional areas sensitive to visual motion lie within the STS, including the medial superior temporal (MST) area and the floor of the STS (FST) (Desimone and Ungerleider, 1986). MST shares a number of properties with MT, such as a high proportion of direction selective neurons and link between activity in this area and the perception of motion (Celebrini and Newsome, 1994). The retinotopic organization in MST is much coarser than that seen in MT, and cells in this area tend to have much larger RFs that often extend into the ipsilateral visual hemifield (Desimone and Ungerleider, 1986). Additional functional differences also exist between the two areas, with many MST neurons shown to be selective for certain complex optic flow motions such as expansion, contraction and rotation, making them candidates for the processing of self-motion (Saito et al., 1986).

The human homologue of V5/MT was first localised by the finding of patients with akinetopsia ('motion blindness') after brain damage to temperoparietal areas, suggesting motion specialized areas in human visual cortex (Zihl et al., 1983). Subsequent fMRI studies found a strong response to moving compared to stationary stimuli around this area (Watson et al., 1993) , suggesting that it is the homologue of macaque area MT. The area of brain activated by motion is also believed to contain the homologues of a number of motion sensitive areas beyond MT (such as MST), so is referred to as MT+. The human homologue of MT (hMT) itself can be identified as the subsection of MT+ that contains a retinotopic map for the contralateral hemifield, and that responds only to motion in the contralateral hemifield (Huk et al., 2002), with an adjacent area showing ipsilateral responses designated as the human homologue of MST (hMST). Axis of motion columns, but not direction of motion columns, have also been demonstrated using high-resolution fMRI in hMT (Zimmermann et al., 2011).

1.2 Magnetic Resonance Imaging

Functional Magnetic Resonance Imaging (fMRI) is non-invasive method of measuring neural activity, with a high spatial resolution. fMRI can be considered a 4D version of anatomical MRI, which can form an image of the brain based on the properties of the underlying tissue when placed in a strong magnetic field. The following section provides an overview of the physics of generating an MR image, what these images tell us about the brain, and how 4D fMRI images (or a timeseries of 3D images) can be related to brain activity.

1.2.1 Nuclear Magnetic Resonance

Magnetic Resonance Imaging (and its functional form), relies on the properties of nuclei and their interactions with magnetic fields. Nuclear Magnetic Resonance (NMR) is the study of nuclei under the influence of a magnetic field. Certain kinds of atomic nuclei will behave in a predictable way when placed within a magnetic field, and this behaviour, and especially their behaviour when perturbed by an electromagnetic pulse, can tell us about the underlying tissue those atomic nuclei are a part of.

Nuclei and Spins

In order to be studied with NMR, atomic nuclei must have a non-zero 'spin', an intrinsic quantum property determined by spin quantum number S. Nuclei with an odd number of protons/neutrons, for instance Hydrogen atoms consisting of a single proton, have a non-zero S. This property gives these nuclei an angular momentum ($\mathbf{J}$), and if a nucleus has spin then it will also possess a magnetic moment ($\boldsymbol{\mu}$) due to the inherent charge of the nucleus. Both of these forces can be represented as vectors with the same direction, related by the scalar factor γ, the gyromagnetic ratio:

$$\boldsymbol{\mu} = \gamma.\mathbf{J} \tag{1.2.1}$$

This scalar factor γ is the ratio between the charge and mass of a given spin. The gyromagnetic ratio is constant for a given nucleus of an isotope with spin, which is key in allowing the development of Magnetic Resonance images. The most abundant and biologically relevant example of a nucleus with spin is hydrogen, which due to its abundance in the human body is most commonly studied in MRI. Hence hydrogen nuclei will be used in subsequent discussion of spins.

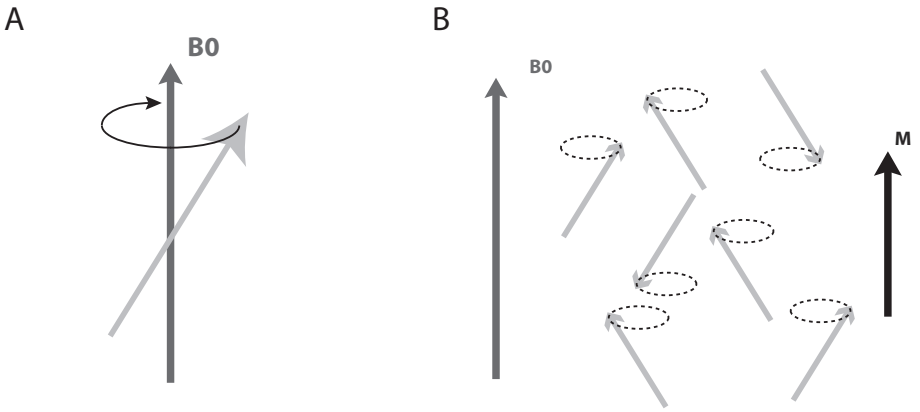

Figure 1.5: Atomic nuclei in a magnetic field. Left panel shows a spin in a parallel state, precessing around B_0. Right panel shows a number of spins aligned either parallel or anti-parallel with B_0. At equilibrium, a greater proportion are aligned parallel with B_0, creating the net magnetization vector **M** parallel with the external field.

If a large number of spins share a spatial location, their magnetic moments sum together to form a net magnetization vector (**M**). In the absence of a strong external magnetic field, the orientations of the axes of the individual spins will be distributed randomly, so will cancel each other out and lead to a very small **M**. However, when placed in a strong magnetic field (known as $\mathbf{B_0}$ in the case of an MRI scanner), the $\boldsymbol{\mu}$ of each spin will experience a turning force trying to align it to the external magnetic field (Figure 1.5). In a steady state the spins do not align exactly with the external magnetic field, but due to their own inherent spin precess around an axis aligned to the external field in a gyroscopic motion. The frequency with which they precess around this axis is defined by the Larmor equation:

$$\boldsymbol{\omega} = \gamma.\mathbf{B} \tag{1.2.2}$$

The frequency of precession is therefore know as the Larmor frequency ($\boldsymbol{\omega}$), and is unique for each isotope due to its relation to the gyromagnetic ratio.

Because of the spin quality of each system, the spins will align either parallel or anti-parallel with the external field (Figure 1.5). The two alignment states have different energy states, low for parallel and high for anti-parallel, and it requires application of energy to cause a transition from the low- to high-energy state. This energy can be provided by the environment a spin is in, leading to a mix of up and down states in a population of spins. Generally there will be more spins in the parallel (low-energy)

state than the anti-parallel (high-energy) state, yielding a small (**M**) vector parallel with $\mathbf{B_0}$. This magnetisation represents the signal available for NMR. The more spins there are in parallel, the larger **M** will be. The proportion of parallel spins, and hence **M**, can be increased by either decreasing the temperature of the spin system (removing the energy from the environment), or increasing the strength of the magnetic field. The net magnetization **M** can be thought of as a vector with two components, a longitudinal component M_z that is parallel/anti-parallel to $\mathbf{B_0}$, and a transverse component M_{xy} that is perpendicular to the main magnetic field.

Radiofrequency Pulse

The net magnetisation vector **M** is not directly measurable under equilibrium conditions because of the huge difference in strength between it and the external magnetic field vector $\mathbf{B_0}$ ($\mathbf{B_0} \gg \mathbf{M}$). To overcome this the equilibrium state is perturbed, which is achieved by applying a radio-frequency (RF) pulse at the Larmor frequency ($\boldsymbol{\omega}$) of the spins. The RF pulse can be thought of as an oscillating magnetic field ($\mathbf{B_1}$), and although it is much smaller in magnitude than $\mathbf{B_0}$, it can change the state of the spins in **M** if it is resonant with the precession frequency of those spins. The application of $\mathbf{B_1}$ causes the the proportion of spins in the low and high energy states to change from that at equilibrium, with more spins entering the high energy state. The application of $\mathbf{B_1}$ also causes the phases of the spins to align. This can be thought of as 'tipping' the magnetic vector from its alignment with $\mathbf{B_0}$ along the z-axis. The actual motion of **M** is a complex spiral motion that is a combination of tipping into the xy plane and precessing around the z-axis, known as nutation. Because both $\mathbf{B_1}$ and **M** rotate around the z-axis at the Larmor frequency, it is often simpler to conceive the action of the $\mathbf{B_1}$ in a rotating frame of reference, which is achieved by changing the coordinate system from the 'laboratory frame' (x, y, z) to a frame of reference rotating at the Larmor frequency known as the 'rotating frame' (x', y', z') Figure (1.6). The z and z' axis are the same in both coordinate systems. In the rotating frame, $\mathbf{B_1}$ appears as a stationary vector along x', and **M** can be considered as simply 'tipping' from the z' axis towards the xy' plane. An RF pulse is described by the angle it 'flips' the **M** vector from the z' axis towards the xy' plane, known as the flip-angle α. The flip-angle is determined by the length of time that the RF pulse is applied. For example a 90° RF pulse will tip the **M** entirely into the $\mathbf{xy'}$ plane, by causing all the spins to be in phase with no net difference between the number of spins in the 'up' and down' state, leading to a net magnetization in the xy' plane. A spin perturbed from equilibrium by a 90° RF pulse is said to be 'saturated' or 'excited'.

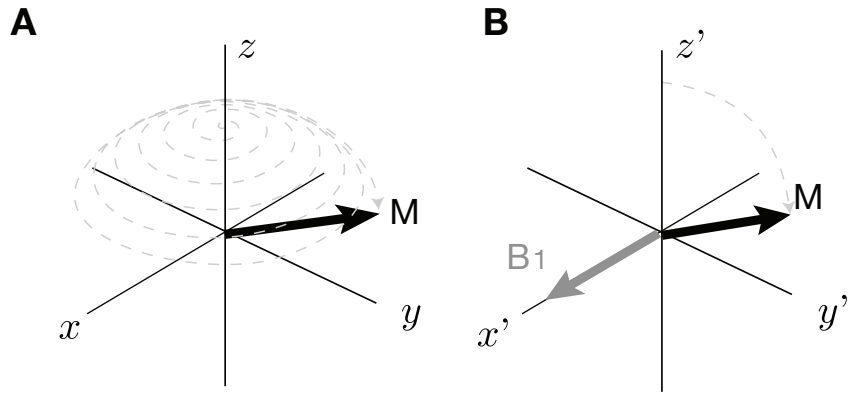

Figure 1.6: The effect of the RF pulse in the laboratory frame (A) and the rotating frame (B). A) In the laboratory frame, the application of the RF pulse tips the magnetization vector **M** into the xy plane as it spirals around the z axis (nutation). B) In the rotating frame, **M** remains stationary, and simply tips into the xy plane.

Relaxation

The result of a 90° RF pulse is a net magnetisation precessing in the xy' plane (due to the net-phase caused by the phase coherence of the spins in the system), and non-equilibrium magnetisation in the z' direction caused by a change in the number of spins in the up and down states. Once the RF pulse is switched off the nuclei that contribute to **M** return to their initial equilibrium state, with a net magnetisation in the z' direction and no coherence amongst the spins, through a process known as relaxation. During relaxation, **M** returns to alignment with z' from being tipped into or through the xy' plane, and the two component vectors return to their equilibrium state. During relaxation, the transverse M_{xy} component decays to 0, and the longitudinal M_z component returns to its equilibrium value. The rates of relaxation for the two components are determined by two time constants, which differ from tissue to tissue.

Longitudinal Relaxation (T_1)

Following the termination of the RF pulse, the spins placed in the anti-parallel, high energy state return to the low-energy state, and the proportion of high and low energy spins returns to the equilibrium point. During this process of longitudinal relaxation, the M_z component returns back to the equilibrium state. For example, following a 90° pulse that tips **M** fully into the transverse plane, the M_z component will be 0. As **M** returns back to alignment with z-direction, M_z recovers from 0 back to its original

value at equilibrium. The recovery of the M_z vector back to equilibrium is governed by a process called spin-lattice relaxation, the exchange of energy between an excited spin and its surroundings. The speed of this recovery, on the order of 1000 ms (depending on the strength of $\mathbf{B_0}$), is defined by the time constant T_1.

Transverse Relaxation (T_2, T_2*)

The application of the RF pulse causes the phases of the spins in **M** to align and tips the magnetization vector into the xy plane, creating the transverse magnetisation vector M_{xy}. When the RF pulse is switched off, interactions between the spins will cause a loss of coherence in the transverse magnetisation, leading to M_{xy} decaying exponentially to 0. This process is sometimes called spin-spin relaxation, and the timecourse of this decay is defined by the time constant T_2. T_2 is much shorter than T_1, ranging from around 10-200ms. The value T_2 is less dependent on the strength of $\mathbf{B_0}$ than T_1.

In practice, the decay in the transverse magnetization M_{xy} is more rapid than expected from T_2, due to inhomogeneities in the magnetic field $\mathbf{B_0}$. These inhomogeneities cause spins in different parts of $\mathbf{B_0}$ to precess at slightly different frequencies, leading to an additional loss of coherence. The resultant relaxation time constant is known as T_2*. Although steps can be taken to correct for this extra decay to measure the true T_2, fMRI takes advantage of the T_2* decay. Specifically, it takes advantage of the fact that T_2* is affected by amount of oxygen in the blood.

Reading the Signal

The change in transverse magnetisation can be detected using the RF coil used to apply the RF pulse, and is the basis of MR signal. The measurable MR signal is proportional to the transverse component, so a larger transverse M_{xy} component will yield a larger signal. Therefore a 90° RF pulse is often used, as this tips **M** fully into the xy' plane, and maximises the signal. The MR signal measured by the RF coil is an oscillating wave at the Larmor frequency, that decays exponentially after the termination of the RF pulse, in a process known as Free Induction Decay (FID) (Figure 1.7). The exponential decay envelope of the signal is defined by the T_2* parameter (Figure 1.7, thick black line), which is a combination of the phase differences between spins caused by spin-spin interactions (T_2decay) and the phase differences caused by spins precessing at different frequencies due to local magnetic field inhomogeneities. If a refocussing pulse is applied at time $t = \tau$ (which rotates the transverse magnetization by 180°) the difference in precession frequencies will cause the spin phases to re-cohere, creating an increase

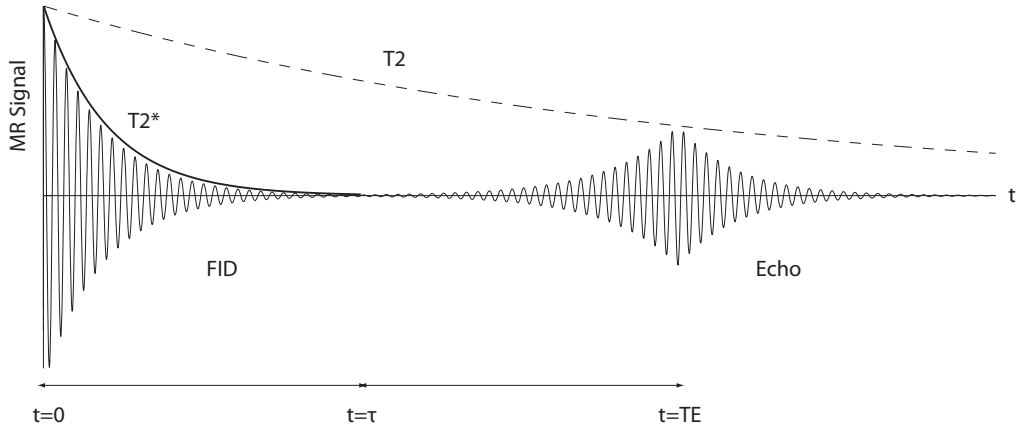

Figure 1.7: MR Signal decay and echo formation. The transverse magnetization component generates a Free Induction Decay (FID) nuclear magnetic resonance signal in the receiver coil, which shows rapid decay defined by T_2*(thick line). After a refocussing pulse at t=τ, a spin echo forms at t=TE, with the peak amplitude of the echo defined by T_2decay (dashed line).

in intensity or echo with a peak at time $t = 2\tau$, with the peak amplitude defined by the T_2 decay parameter (Figure 1.7, dashed line). This is the basis of *spin-echo* (SE) imaging, which is one of the methods used to create MR images. An alternative method of echo formation uses the magnetic gradients utilised in 2D image formation to create the echo, and is known as *gradient-echo* (GE) imaging. This method does not negate the phase differences caused by spins precessing at different frequencies, so the peak amplitude of the echo is defined by the T_2*decay parameter rather than T_2. In both cases, the time between the RF excitation pulse and the peak in the echo is referred to as the echo time (TE). The time between repetitions of the RF excitation pulse is referred to as repetition time (TR).

The choice of TR and TE for a given pulse sequence will emphasise different aspects of the signal, specifically the contrast in signal strength recorded from tissues with different T_1 and T_2 values. By choosing a very short TR and TE, which does not allow the longitudinal magnetization to return to equilibrium between RF pulses, the strength of the recorded signal is primarily defined by T_1 value of the underlying tissues. Images collected with these TR and TE values are known as T_1-weighted images. Sequences with a longer TR, where longitudinal magnetization returns to equilibrium between RF pulses, emphasise differences in T_2/T_2*, and lead to T_2/T_2*-weighted images, depending on the use of a refocussing pulse.

BOLD FMRI (discussed more fully in section 1.2.3) relies on the fact that T_2* varies

according the amount of deoxygenated blood present in a given cortical area, and so tends to utilize GE sequences. The choice of TE can be a deciding factor in defining BOLD contrast, and will be maximal when the TE used for echo formation matches the T_2* of the underlying gray matter.

1.2.2 Magnetic Resonance Imaging

The conversion of the NMR signal to a 3D MRI image, by the addition of information about the spatial origin of the signal from the spins, requires the application of magnetic gradients at different stages of the signal generation and readout process.

Slice Selection

A gradient applied in the z-direction G_z at the same time as the RF pulse will cause only a restricted 'slice' of the spins in the object along the z-axis to become excited, and this gradient is known as the 'slice selection' gradient (Garroway et al., 1974).

Phase/Frequency Encoding

The application of a gradient in the y direction (G_y) whilst the net magnetization is in the transverse plane will create a linear spatial variation in the phase of the transverse magnetization, varying along the y-axis. A gradient in the x-direction applied during image readout (G_x) creates linear spatial variation in the precession frequency of the spins along the x-axis. Therefore, the spatial position of the signal in the z-direction is defined by which areas are excited by the RF pulse, and x and y locations are encoded by the frequency and the phase components of the recorded signal.

If a refocussing pulse is used (SE imaging), the G_x gradient is applied during the formation of the echo. Alternatively the gradient itself can be used to create an echo (GE imaging) (Figure 1.8). In this method, an RF pulse is applied (Top Row), with a G_z gradient for slice selection (Second Row). A phase encoding gradient G_y is applied next (Third Row), along with a dephasing frequency encoding gradient (Bottom Row). This gradient is negative in sign from that of the frequency encoding gradient G_x which is turned on during the acquisition of the signal. An echo is produced (Bottom Row) when the frequency encoding gradient is turned on because this gradient refocuses the dephasing which occurred from the dephasing gradient.

To generate a 2D MR image from the excited slice of tissue requires full mapping of 'k-space' for that image. K-space represents the spatial frequency distribution of the

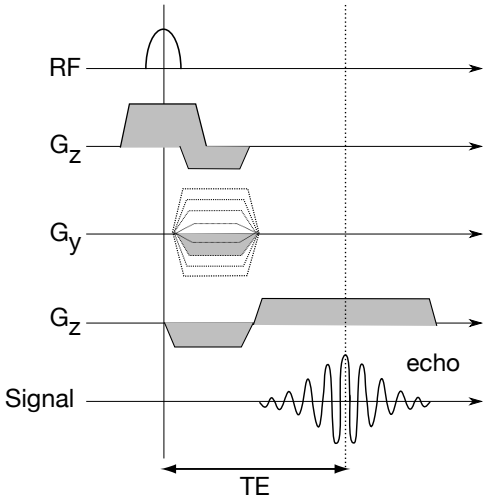

Figure 1.8: Example of a gradient echo (GE) imaging pulse sequence. The top row shows the the RF pulse. The second row shows the G_z gradient applied with the RF pulse. The third row shows the G_y gradient, and the fourth row shows the negative and positive lobes the of the G_x gradient. The positive lobe of the G_x gradient causes a gradient echo to form (Bottom Row), which peaks in amplitude at time TE.

image, with low spatial frequencies lying at the centre of k-space, and fine detail information appearing towards the edges. The gradients used in encoding an image after an RF pulse define how k-space is sampled. Once k-space has been sufficiently filled, an inverse Fourier transformation will convert the data from k-space to conventional image space. The image reconstruction process yields an image of the activated slice made up of a series of *voxels* (the 3D equivalent of pixels in an image), with the resolution of the image determined by the sampling of k-space.

In standard imaging, for example that used to acquire anatomical imaging, the full sampling of space is done piecemeal (generally one 'line' of k-space at a time), requiring multiple RF excitations to acquire a full image. Mansfield (1977) developed echo-planar imaging (EPI), a fast imaging modality that mapped the entire k-space after application of a single RF pulse, rather than using multiple RF excitations to sample k-space. The use of EPI allows images to be acquired rapidly enough to study the changes in blood oxygenation that result from neural activity (functional MRI). However, the methods required to acquire images quickly make these images particularly susceptible to distortions and artefacts caused by non-uniform magnetic fields, and these issues can be especially prevalent at higher magnetic field strengths.

The sequence described above is 2D imaging, where multiple 2D slices are collected to

be combined into a final, 3D image. 3D MR imaging is also possible. In 3D imaging, a large slab (as opposed to a thinner slice) of tissue is excited by the RF pulse, and localization in the z-direction is done with an additional phase encoding step. K- and image-space are now three-dimensional, and an image is formed from the data in k-space via a 3D inverse Fourier transformation. In general, anatomical imaging is often done in 3D whilst functional imaging is done in 2D, although recently 3D imaging has been used for high-resolution fMRI at 7T.

Recent developments have allowed stronger magnets to be used in MRI scanners, with field strengths of 7T and even 9T now available for use with humans. One key benefit of using high-field strengths is the increase in image signal to noise ratio (SNR), which allows MR images with greater resolution (smaller voxel sizes) to be obtained. However, the benefits afforded by using high-field imaging come with a number of technical challenges compared to standard-field imaging.

Issues with MR Imaging

When collecting MR images, inhomogeneities in the static magnetic field (B0) can cause artefacts in the images, manifesting as either loss of signal in certain areas (drop-out) or distortions in the images. As these issues (known as susceptibility artefacts) scale with magnetic field strength, additional steps are required in high-field imaging to correct them.

Signal loss is caused by dephasing of spins within a voxel due to small-scale magnetic field inhomogeneities, leading to a more rapid decay in signal due to T_2* effects. Effects such as these are particularly strong at the border between tissue and air such as the sinuses and ear canals, and mean that imaging in areas such the medial-frontal and ventral-temporal cortices can be difficult. In addition, signal drop-out in these areas increases at higher-field strengths (Poser and Norris, 2009). Potential solutions to this issue include shortening the TE for imaging to match the $T_2*$of the drop-out areas (which can however decrease sensitivity elsewhere in the brain) or the use of 'double-echo' imaging methods (Poser et al., 2006). As our experiments were primarily focussed on visual cortex, where signal drop-out is less of an issue, no specialized methods were required to deal with drop-out.

Variations in magnetic field strength will also lead to variations in the spin frequency of the underlying spins, leading to inaccurate localization of those spins due to frequency-encoding. This manifests as distortions in the image. These distortions become especially apparent during rapid image acquisitions, for example EPI, as the errors will

accumulate over the relatively long read-out time for these methods. Inhomogeneities in the magnetic field are also more severe at higher field strengths, therefore distortions in EPI at 7T can be especially problematic. Methods to combat these distortions include parallel imaging techniques such as SENSE (using multiple coils to read the signal, which can reduce the number of phase-encoding steps to speed up imaging whilst allowing the same k-space sampling) and accounting for and correcting the inhomogeneities in the magnetic field. If the magnetic field itself can be mapped, these residual inhomogeneities can be corrected, leading to less distortions in the collected images.

1.2.3 Functional MRI

As well as allowing for anatomical images of the brain to be collected at high-resolutions, MR imaging can also give us indirect measures of neural activity. The relaxation rates of different tissues vary in a reliable way based on biological processes related to the underlying neural activity, and these changes can be measured using MR imaging methods with sufficient temporal resolution such as EPI.

The BOLD Signal

The magnetic properties of haemoglobin in the blood differ depending on the presence/absence of oxygen. Oxygenated haemoglobin is diamagnetic, whereas deoxygenated haemoglobin is paramagnetic, leading to greater magnetic susceptibility of deoxygenated blood. Deoxygenated blood causes nearby spins to precess at different frequencies, leading to destructive interference, and a shorter T2*. Therefore, blood with a greater proportion of oxygenated blood should lead to a larger MR signal than deoxygenated blood in T_2* weighted images. This effect was first demonstrated in rats, where it was found that scanning the brains of rats breathing normal air (21% oxygen) yielded different GE images compared to rats breathing 100% oxygen. The scans from the 100% oxygen rats showed standard contrast between tissue types. The scans from the 21% oxygen rats showed dark lines in areas corresponding to blood vessels, and if the rats instead breathed a gas mixture with 0% oxygen, the lines became even darker. This change in MR signal based on blood oxygenation is known as the Blood Oxygen Level Dependent (BOLD) contrast (Ogawa et al., 1990), and is the basis of functional MRI.

It may be expected that an increase in neural activity should lead to a decrease in BOLD signal, due to increased oxygen consumption from the metabolic demands. However,

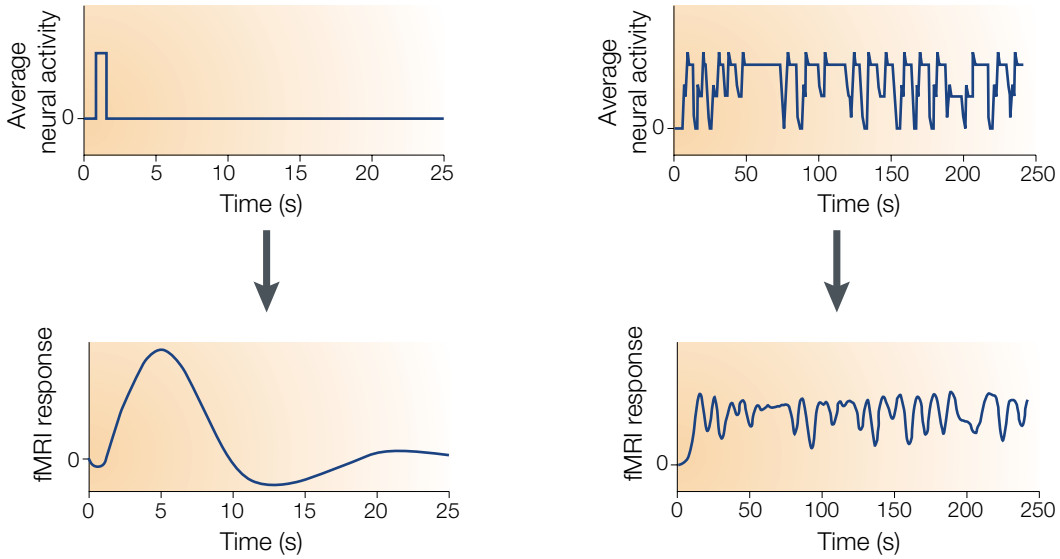

Figure 1.9: The hemodynamic response. Left column shows a 'canonical' HRF in response to a brief burst of neural activity, showing the initial dip, the delayed peak and the undershoot. Right column shows how the HRF is expected to sum for an extended series of peaks in neural activity. Figure taken from Heeger and Ress (2002).

following neural activity, a paradoxical increase in BOLD signal is in fact found. The reason for this is that even though the neural activity leads to an initial increase in oxygen consumption and a decrease in blood oxygen levels, this is followed by an increase in cerebral blood flow (CBF) and a subsequent oversupply of oxygenated blood to that area of cortex, causing an increase in BOLD signal (see Heeger and Ress (2002) for a review). The BOLD signal is based on the interplay of oxygen consumption, CBF and blood volume. BOLD contrast increases with field-strength (Yacoub et al., 2001b), leading to an increase in fMRI studies at higher field-strengths.

The Hemodynamic Response

Due to the indirect nature of the signal generation, the change in BOLD signal in response to neural activity has a distinctive timecourse, known as the Haemodynamic Response Function (HRF) as shown in Figure 1.9. The HRF has 3 distinct features, relating to the distinct aspects of the blood oxygenation in response to neural activity. In the first second or so after a burst of neural activity there is a drop in signal, although it can often be fairly subtle and it not always observable. This is believed to relate to the initial increase in deoxygenated blood following oxygen consumption related to neural activity. Subsequently there is a larger increase in signal, peaking ~6 seconds after the

neural activity, due to the subsequent increase in CBF and oversupply of oxygenated blood. It has been suggested that the initial dip may be more tightly correlated in space with neural activity than the influx of oxygenated blood, but generally the *increase* in signal is easier to measure. Following the peak in signal there is a decrease back towards baseline, an undershoot where the response is actually negative, and then a final return to baseline. This undershoot has been related to a number of physical properties both in terms of blood flow/volume (Buxton et al., 2005) and metabolic effects (van Zijl et al., 2012).

An assumption that the HRF is proportional to the underlying neural activity allows the BOLD timecourse in response to stimuli to be modelled. This assumption is known as the linear transform model (Boynton et al., 1996), which assumes that the same HRF is evoked by a stimulus independently of how close in time it is presented to another stimulus, and that the HRF of stimuli presented close in time will sum together (superposition). Evidence for a 'rough' linearity has been shown (Boynton et al., 1996), although a stimulus presented very soon after another (during a 'refractory' period) will show a reduced response. This particular non-linearity can be exploited in experimental techniques such as fMRI adaptation. In general though, the BOLD signal in response to a sustained or repeated stimulus presentation is modelled by convolving a model HRF (the BOLD response to a brief stimulus) with the stimulus pattern. The amount of variance in the actual timeseries explained by this modelled timeseries can then be assessed. Full details of the process used to model the BOLD timeseries are given in Chapter 2.

Limitations and Improvements of BOLD fMRI

Although the development of fMRI has allowed the non-invasive study of neural activity, there are a number of potential limitations to this method, stemming from the fact that the BOLD signal is a signal of hemodynamic origin with an indirect link to underlying neural activity. Some of these issues can be mitigated or avoided by specialized techniques or improvements in scanner technology. Others are inherent to the technique itself and cannot be avoided.

A major consideration when using fMRI to study neural activity is the indirect nature of the BOLD signal. The exact link between neural activity and changes in BOLD signal is not fully understood, especially whether the coupling between the two is loose or tight. Combined recordings of neural activity , both in terms of neuron spiking and local field potentials (LFP), and BOLD in monkeys show that BOLD signal is more closely correlated with LFP than the spiking activity. LFPs are related both to post-

neuron synaptic activity and intra-neuronal processing, reflecting both the input to a neuron and its internal activity. This means that both excitation and inhibition of the cell can be reflected in the LFP, whereas an increase in spiking can only reflect excitation (Logothetis, 2008). Therefore drawing conclusion about neural activity can be non-straightforward based on an increase or decrease in the BOLD signal.

Because of the hemodynamic nature of the BOLD signal, the spatial resolution of the BOLD signal may also be inherently limited. The BOLD signal in response to neural activity in a given area will reflect the contribution both from capillary beds that are tightly localized to the activity in question, and draining veins that may be spatially distant from the cortical area in question (Frahm et al., 1994).

Two possible methods of increasing the spatial specificity of the BOLD signal exist: increasing the magnetic field strength and using spin-echo rather than gradient-echo imaging. Increasing the magnetic field emphasises the signal coming from spins in the tissue affected by the field inhomogeneities caused by de-oxygenated blood (the extravascular signal) over the signal from spins within the blood vessels themselves (the intravascular signal). This suppression of the intravascular (IV) signal is due to the shortening of T_2* for venous blood compared to tissue at higher fields (Yacoub et al., 2001b), leading to the IV component of the signal decaying away more rapidly than the extravascular (EV) component. Thus the IV signal from large draining veins is suppressed at higher field strengths, increasing spatial specificity. However, for some imaging techniques such as GE imaging, the EV component of the signal can arise from both large and small vessels, which can reduce the spatial specificity of the signal. Using SE rather than GE imaging can suppress the EV signal from larger veins, leaving only the signal from the smaller vessels, although the sensitivity to the BOLD signal is reduced. Combing high-field imaging with SE image acquisition is thought to provide the most spatially specific signal (Yacoub et al., 2003), although at the cost of signal amplitude that precludes the use of SE imaging in most instances.

It has also been suggested that the initial dip is more closely related to neural activity, representing the initial increase in de-oxygenated blood due to oxygen extraction in the capillaries after neural activity, whereas the subsequent peak represents an over-compensatory increase in oxygenated blood that is less spatially restricted (Menon et al., 1995). However, capturing the initial dip can be problematic, due to both its transient nature and due to the fact that it has a very small amplitude at lower fields, consistent with the suggestion it arises from effects in smaller blood vessels. This presents the possibility that it may be more apparent at higher field strengths (Yacoub et al., 2001a).

At higher field-strengths, both image SNR and BOLD signal magnitude are increased. Furthermore, an increase in field-strength can also allow an increase in resolution, which can allow an increase in BOLD contrast due to the reduction in partial voluming effects, where voxels contain a mixture of grey matter, white matter and CSF. Small voxels can also increase the functional contrast of the signal by isolating tissue with a uniform selectivity. However, in both cases there is a trade-off between voxel-size and SNR, which can set a minimum threshold for voxel size.

1.3 Experimental techniques

Over the years, a number of different analysis techniques have been developed for fMRI datasets. These include examining the change in amplitude between two blocks of task or stimuli or evaluating the change in the amplitude or shape of the HRF between trials of a given task or stimulus. One key similarity between these various techniques is that they either treat the timeseries from each voxel as separate from those around it during analysis, or alternatively average the timeseries from a number of voxels together prior to analysis. This approach to fMRI analysis can be described as *univariate* because only one variable, the timeseries from the voxels or ROI, is considered in the analysis at a time. Univariate analyses either yield statistical maps of activity across the brain, or a single value per ROI. One potential limitation of this approach for measuring neural activity is that the signal measured from a given voxel/ROI can include contributions from a number of cells with a range of different selectivities, which can make distinguishing between responses to certain stimuli (for example those for different orientations or directions of motion) problematic.

1.3.1 Adaptation

To investigate the functional properties of sub-voxel neural populations, without requiring very-fine spatial resolution images, a number of alternative methods have been developed. The first is adaptation, which has a long history of use in psychophysical and neurophysiological experiments. Exposure to a visual stimulus will affect the neural responses to, and the perception of, subsequently viewed visual stimuli. The nature of these changes can provide us with evidence about the neural mechanisms involved in visual perception.

Perceptual adaptation exists in many domains, and motion stimuli provided one of the first examples of perceptual adaptation effects: extended viewing of a moving stimulus causes subsequently viewed stationary or ambiguous stimuli to be perceived as moving in the opposite direction (the motion after-effect) (MAE) (Thompson, 1880). Similarly, the apparent contrast of a grating displayed after prolonged exposure to a grating with similar properties will be reduced, and the reduction in apparent contrast depends on the similarity between the adapter and the probe (Blakemore and Campbell, 1969). Furthermore, the perceived orientation of a probe grating can be affected by extended exposure to another oriented grating. Adaptation effects for the detection and discrimination of direction of motion have also been found, with the strength of the effect depending on the similarity between adapter and probe (Hol and Treue, 2001).

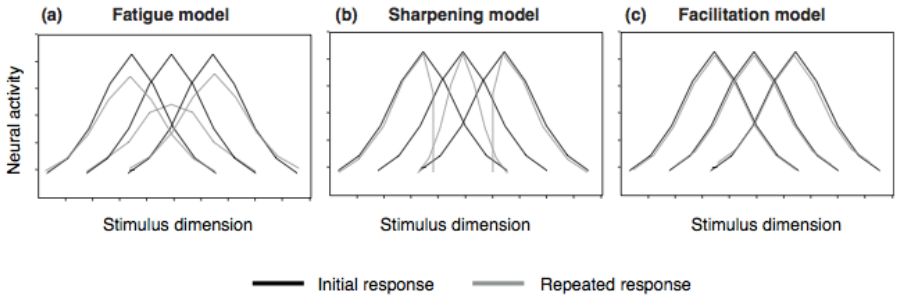

Figure 1.10: Pre (Black) and post-adaptation (Grey) tuning curves for three exemplar neurons under three models of adaptation. The adapting stimulus in each case is the central stimulus. A) 'Fatigue' model: responses are reduced after adaptation in proportion to initial response to adapting stimulus, with no change in tuning width. B) 'Sharpening' Model: adaptation narrows the tuning curves around the adapting stimulus. C) 'Facilitation' model: no predictions are made for changes in the tuning curves. Figure taken from Grill-Spector et al. (2006).

More recently, adaptation to higher-level stimulus properties, such as the perceived gender of faces, has also been demonstrated (Webster et al., 2004).

Perceptual after-effects and changes in the detection and discrimination of stimuli provide evidence of selective neural mechanisms in the human brain. Psychophysical results can provide us with evidence as to the neural coding underlying perception, providing some of the first evidence of orientation and direction-selective neurons in the human visual system.

The neural consequence of adaptation are not fully understood at this point. One popular model of the neural effects of adaptation is the so-called 'fatigue model' (Figure 1.10, A). In this model, neurons selective for the adapting stimulus reduce their firing rates after extended or repeated exposure to that stimulus. This leads to a general reduction in firing rates across the population of cells, with the greatest reduction seen in the cells that respond most strongly to the adapting stimulus. Such a change in selective properties would have many potential benefits, such as increasing sensitivity for novel stimuli or optimising the dynamic range of the selective neurons. Measuring the tuning curves of cells after adaptation has demonstrated the fatigue model is sufficient to explain adaptation affects for orientation and motion in V1 cells (Kohn and Movshon, 2004).

A second model suggests that adaptation instead leads to a narrowing ('sharpening') of the tuning curves for selective neurons, leading to fewer neurons responding to subsequent presentations of that stimulus (Figure 1.10, B). Under this model, neurons

optimally selective for the adapting stimulus would see the least reduction in activity, rather than the greatest levels. Evidence shows that this model of adaptation may explain the effects of adaptation seen in MT for motion, contrasting with the 'fatigue' like effects seen in V1 (Kohn and Movshon, 2004). A third model posits that tuning is unaffected by adaptation, instead suggesting that the processing of stimuli is made quicker ('facilitation') (Figure 1.10, C). Different models of adaptation may apply in different brain areas and under different task/stimulus conditions (see Grill-Spector et al. (2006) for a review).

Selectivity along a particular stimulus dimension can also be demonstrated in fMRI using the methods of adaptation (Grill-Spector and Malach, 2001). These methods use extended exposure to a stimulus followed by testing with a probe stimulus that either matches or differs from the adapter along the dimension of interest. The BOLD response to an identical stimulus will be reduced in areas sensitive to that stimulus. By measuring the response to stimuli that differ from the adapting stimulus, and seeing whether this reduction in BOLD response is stimulus selective, then neural selectivity for the dimension in which the stimuli differ can be inferred.

Measuring the effects of fMRI-adaptation could potentially allow subpopulations of neurons with different selectivities to be differentiated, despite them being co-localised in a single voxel, and hence offers one potential solution to the limitations of fMRI sampling resolution (Grill-Spector and Malach, 2001). How these stimulus-selective reductions in BOLD signal are related to the neural and perceptual effects of adaptation is not fully understood at this point, so care must be taken when interpreting fMRI-adaptation results in terms of the underlying neural responses.

1.3.2 MVPA

An approach that has recently been introduced is to treat the timeseries from multiple voxels as a single multivariate dataset, and to take the activity of many voxels into account when performing an analysis. One such technique is multivariate/multi-voxel pattern analysis (MVPA), which uses the patterns of activity across a number of voxels to create 'classifiers' that can decide which stimulus or cognitive state led to a novel activity pattern. MVPA methods are particularly exciting as they raise the possibility that signals arising from individual populations of neurons or columns can be accessed at standard fMRI resolutions (Boynton, 2005).

Even in cases where signals at the univariate (single voxel) level do not allow stimuli to be differentiated, if the pattern of activity from a number of voxels is consid-

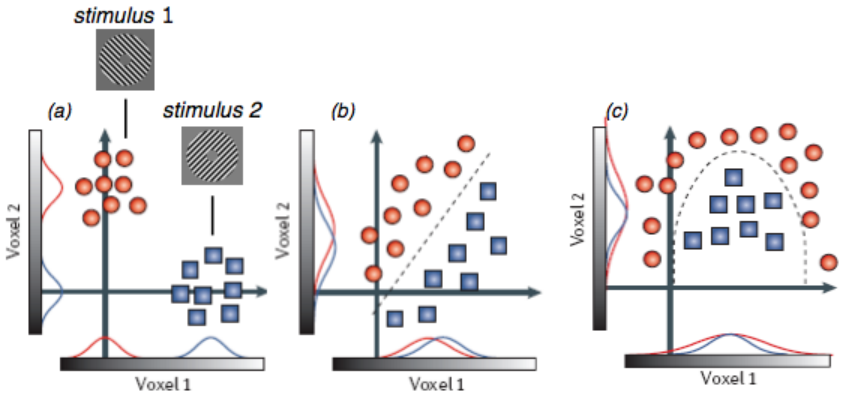

Figure 1.11: 2 Voxel example of how multivariate methods allow classification of stimuli. The left panel shows a situation where the stimuli classes can be identified on the basis of single voxels (univariate). The middle panel shows a situation where the two classes overlap, but a linear decision boundary can be calculated when the 2 voxels are considered simultaneously. The right panel shows where a nonlinear decision boundary is required. Figure taken from Haynes and Rees (2006)

ered multivariately, stimuli can often be differentiated with greater sensitivity. This is demonstrated in Figure 1.11. Each point in the plots represents a trial, with the colour representing a different trial type (in this case orientation stimuli). The left hand panel in Fig. 1.11 indicates a situation where the two stimuli can be differentiated based on the responses of either voxel taken in isolation. The middle and right panels indicate situations where the univariate responses overlap, so the response of either voxel in isolation cannot be used to distinguish between the stimuli. However, in both cases considering the response of both voxels simultaneously allows the two stimuli to be distinguished, by calculating a decision boundary. Any response on one side of this boundary is classified as one stimulus class, anything on the other side is classified as the other stimulus class.

The middle panel shows the linear case, which in the 2-dimensional class yields a boundary line. Activity patterns resulting from a large number of voxels are often used, in which the responses of N voxels are considered in an N-dimensional space. In such an N-d space, the classes are separated linearly by a hyperplane. This method can also be expanded to distinguish more than two classes, in which case multiple hyperplanes are used for each two-way classification. A large number of linear classifiers exist (Duda et al., 2001), including linear support vector machines (SVM) (Schölkopf et al. (1999), see Kamitani and Tong (2005, 2006) for application), linear discriminant analysis

(LDA) (Haynes and Rees, 2005a, 2006), and Gaussian Naive Bayes classifiers (Brouwer and Heeger, 2009). All of these different methods work by computing a weighted sum of voxel activity to calculate a linear decision boundary to decide whether an observed data-point belongs to a particular category.

The right panel in Figure 1.11 represents a non-linear classification, which results in a curved decision boundary. Whilst non-linear classifiers can be more sensitive to the differences between stimuli than linear ones they are sensitive to over-fitting, and may indicate selectivity in areas where none exists. For this reason, linear classifiers are generally used in MVPA experiments (Tong and Pratte, 2012).

The accuracy of the classifier is tested by cross-validation: training on one set of data from a subject, and testing on a separate set from that subject. Often the testing set will be data from a single run, and the training data will be data from the remaining runs, with each run being tested in turn. This procedure avoids training and testing a classifier on the same data. The classification accuracy will generally depend on the separability of the responses to the stimuli (the distance between the groups of points in N-dimensional space), the number of datasets available for training, and the number of voxels used to generate the patterns of analysis. In general, the more voxels are used the more accurate the classifier can be, although if uninformative voxels are added to the patterns they may simply add noise and reduce classifier accuracy (Yamashita et al., 2008).

MVPA methods have been used to decode a wide range of stimulus properties, including orientation (Haynes and Rees, 2005a; Kamitani and Tong, 2005; Swisher et al., 2010), direction of motion (Hong et al., 2011; Kamitani and Tong, 2006; Seymour et al., 2009), eye of origin (Haynes et al., 2005), colour (Brouwer and Heeger, 2009; Seymour et al., 2009), binocular disparity (Ban et al., 2012; Preston et al., 2008), object category (Cox and Savoy, 2003; Op de Beeck, 2010) as well as conjunction of features such as colour and motion (Seymour et al., 2009). In addition MVPA methods have been shown to be able to classify which of two overlayed stimuli were attended to (Kamitani and Tong, 2005, 2006), which of two stimuli were perceptually dominant during binocular rivalry (Haynes and Rees, 2005b), and the perception of ambiguous motion displays (Brouwer and van Ee, 2007; Serences and Boynton, 2007a). Results such as this suggest that an observer's conscious perception can be decoded, and have therefore often been referred to as 'mind-reading', and have sparked a wide range of discussion as to the potential of these kind of techniques, and the potential ethical implications (Tong and Pratte, 2012).

Although the use of multivariate techniques has increased greatly in recent years, it should be noted that the exact mechanism by which multivariate methods make use

of neural information is not fully understood and it has been suggested that caution should be used when interpreting MVPA results (Bartels et al., 2008). One key area of debate is the scale of the signal exploited, specifically whether MVPA methods allow the measurement of sub-voxel level information. For stimuli such as orientation, where neurons are arranged into columns on the cortical surface smaller than the resolution available to typical fMRI techniques, it has been suggested that classification results may be driven by unequal sampling of columns selective for different orientations by the voxels (Boynton, 2005), or local variations in columnar organization (Kamitani and Tong, 2005, 2006). Alternatively, the results may reflect more global, coarse-scale biases (Freeman et al. (2011); Op de Beeck (2010), but see Kamitani and Sawahata (2010); Swisher et al. (2010)). Whilst columnar arrangement of feature selective neurons may allow successful classification, a successful classification result is not necessarily evidence of the presence of columns (Bartels et al., 2008), and high-resolution fMRI (Yacoub et al., 2001b, 2007; Zimmermann et al., 2011) or optical imaging are necessary to provide further proof of such organization.

Chapter 2

General Methods

This chapter summarizes the general shared methods that were used for each experiment in this thesis. This includes the methods to segment and flatten the cortical surfaces, the methods to identify the retinotopic visual areas, the fMRI scanning parameters and the various analysis methods.

2.1 Cortical Segmentation and Flattening

Although the spatial resolution of the EPI images conventionally obtained from fMRI is high in terms of localisation, it is generally not sufficient for BOLD activity to be related to the underlying neural anatomy (although recent high-resolution studies at 7T have showed sufficient resolution to allow this). It is generally the case that a high-resolution anatomical image is obtained separately, onto which the functional data is overlayed. Often these images are T1-weighted due the good contrast between white and grey matter. However the cortical surface is folded and convoluted, so interpreting spatial locations on these 3D images can be difficult. A common method for dealing with this issue is to identify and extract the surface between the grey and white matter, inflate and flatten it, and display the functional data on this flattened representation of cortex. This method allows the spatial locations of functional data to be more easily visualized, and is particularly useful for the visualisation and analysis of retinotopic maps.

The surface-extraction method begins with a T1-weighted anatomical image, which is preprocessed and segmented, and ends with two cortical surfaces per hemisphere, one representing the white-matter (WM)/grey-matter (GM) boundary and the other to the grey-matter/cerebrospinal fluid (CSF) border. The WM/GM boundary can then

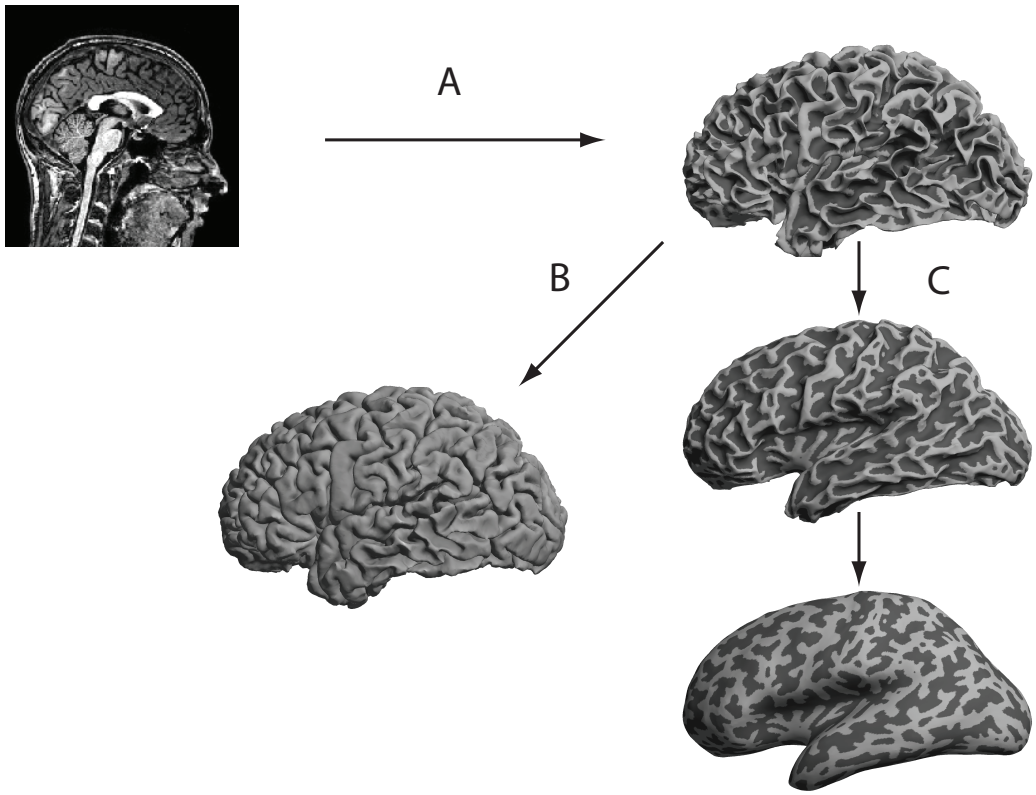

Figure 2.1: Processing stream for cortical surfaces. The white-matter/gray-matter boundary is identified from the T1 weighted anatomy (A), and then expanded to find the gray-matter/CSF boundary (B). The WM/GM boundary can also be inflated to smooth out the cortical sulci and allow easier visualization (C).

be inflated and flattened and used to display the data. The extraction and inflation process for our data was performed using Freesurfer (Dale et al., 1999) (see Figure 2.1 for a description of the process).

The preprocessing begins with registering the T1 anatomy (Figure 2.2 A) into Talaraich space, which is used to generate starting estimates for the location of WM and key areas such as the corpus callosum. This is followed by normalising the image intensities in the T1 image present due to magnetic field inhomogeneities, which could lead to problems in identifying the borders between tissue types, using a non-parametric method (Sled et al., 1998). Non-brain tissues are then removed (for example, skull stripping) (Figure 2.2 B) using a surface deformation procedure (Segonne et al., 2004). Following this, voxels within the image are identified as being WM based on their intensity and Talairach location. The hemispheres are identified and seperated from each other

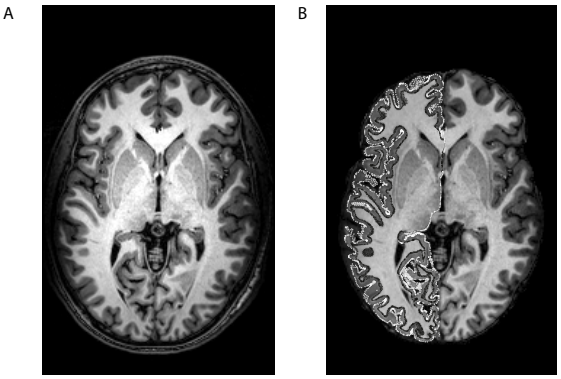

Figure 2.2: Examples of the original T1 weighted 3D image (A) and the preprocessed T1 weighted image with the calculated WM/GM (black) and GM/CSF (white) surfaces overlayed.

and the brain stem, with the hemisphere boundaries based on the likely location of the corpus callosum in Talairach space. An initial surface is then generated for each hemisphere by tessellation of the outside of the white matter mass for that hemisphere, followed by topology fixing. This initial surface is then deformed following intensity gradients to optimally place the gray/white and surface at the location where the greatest shift in intensity defines the transition to the other tissue class. Finally, this surface is then expanded to match the gray/CSF boundary in the same way (Dale et al., 1999) (Figure 2.2 B).

Following the identification and extraction of these surfaces, the GM/WM boundary can be inflated (Fischl et al., 1999). Whilst displaying functional data on either the GM or WM surfaces extracted earlier allows greater appreciation of spatial localisation than display on the 3D anatomy image, the sulci can still hide a large amount of activation. By inflating the surface, the sulci are flattened out and activity can be more easily observed. FreeSurfer inflates the surface while attempting to minimize geometric distortions and maintain the relative geometric relationships in the surface topology. Often the analogy of inflating a crumpled paper bag is used, in that the surface is inflated without any expanding as the surface is not elastic (Pienaar et al., 2008). Once inflated, the curvature values calculated from the original folded surface are typically displayed on the inflated surface. Examples of in inflated surface with the curvature displayed are shown in Figure 2.3. The curvature has been binarized, so that dark areas indicate sulci and light areas indicate gyri. The calcarine sulcus has been indicated on the medial view.

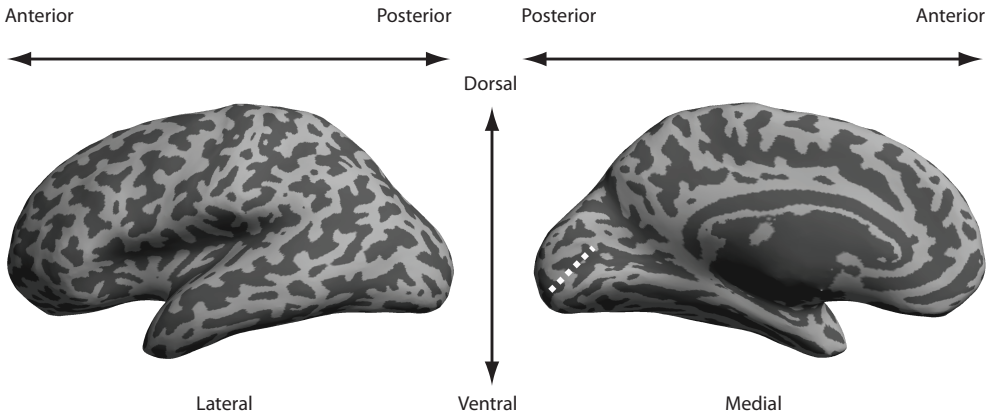

Figure 2.3: Examples of inflated cortical surfaces from the lateral view and medial view. Dark areas indicate sulci, light areas indicate gyri. White dotted line indicates calcarine sulcus.

For the close inspection of data, for example examining retinotopic maps in visual cortex, a further stage of flattening a patch of the inflated cortex can be implemented. We utilised mrFlatMesh, part of custom written software (mrTools, VISTA), to obtain the flattened patches.

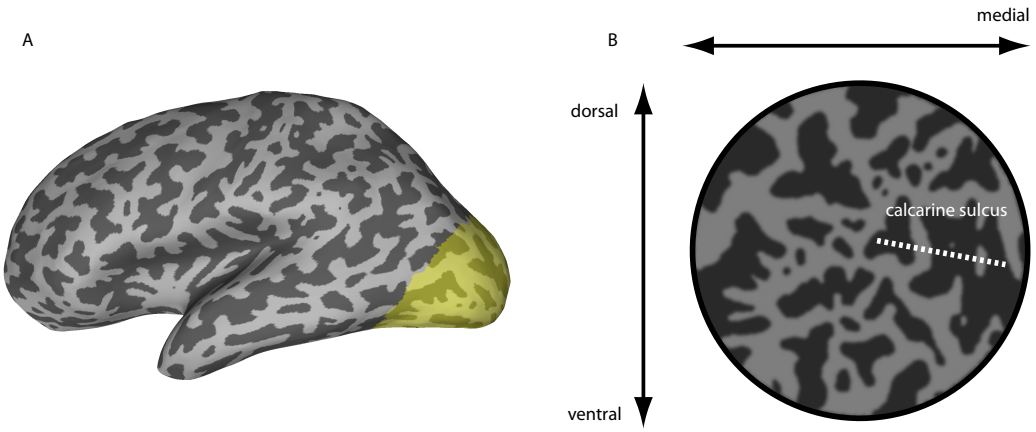

Figure 2.4: A) The location of the flattened patch of cortex on the inflated cortical surface, centered around the occipital pole. B) The flattened patch of cortex, with the calcarine sulcus indicated by the white dotted line.

2.2 Retinotopic Mapping

The visual cortex is retinotopically organized, in that neurons with receptive fields covering adjacent points of the visual field are adjacent on the cortical surface. This organization, called retinotopy, allows the identification of early visual areas. The border between early visual areas (V1, V2, V3 etc.) is demarcated by reversals of the retinotopic map for polar angle. Retinotopic maps to identify these borders can be generated using techniques first described by Engel (Engel et al., 1997). The technique uses stimuli consisting of flashing "checkerboards" that change their position over time to create periodic patterns of activity in retinotopically-organized areas. Examples of these stimuli are shown in Figure 2.5 consisting of shapes formed from a checkerboard pattern, which reverses contrast rapidly to ensure the neurons at that retinotopic location are stimulated.

Maps of eccentricity are measured using expanding and contracting rings of the pattern, whilst mapping for polar angle is measured using a rotating wedge stimulus. Although this specific stimulus set-up was originally used to retinotopically identify the visual areas, a wide range of different stimuli can produce equivalent maps (Dougherty et al., 2003; Dumoulin and Wandell, 2008).

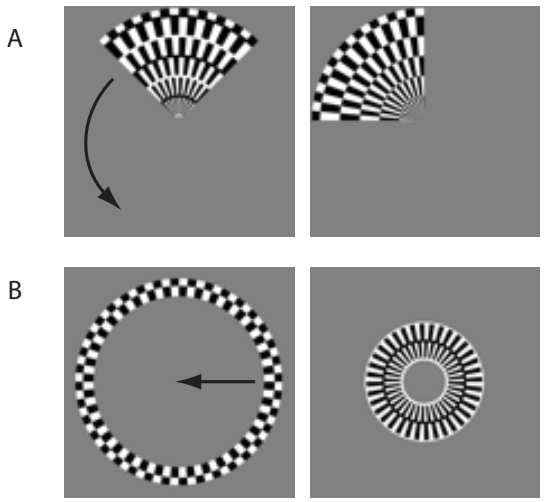

Figure 2.5: Stimuli used in the retinotopic mapping paradigm. A) Rotating wedge stimulus for polar angle mapping. B) Expanding/contracting ring stimulus for eccentricity mapping.

The mapping technique works by making use of the limited receptive fields of neurons in the visual areas. When the retinotopy stimuli pass over a point in the visual field, neurons with receptive fields centred on that area will be stimulated only when the stimulus is at that point. As the stimulus progresses, this leads to a travelling wave

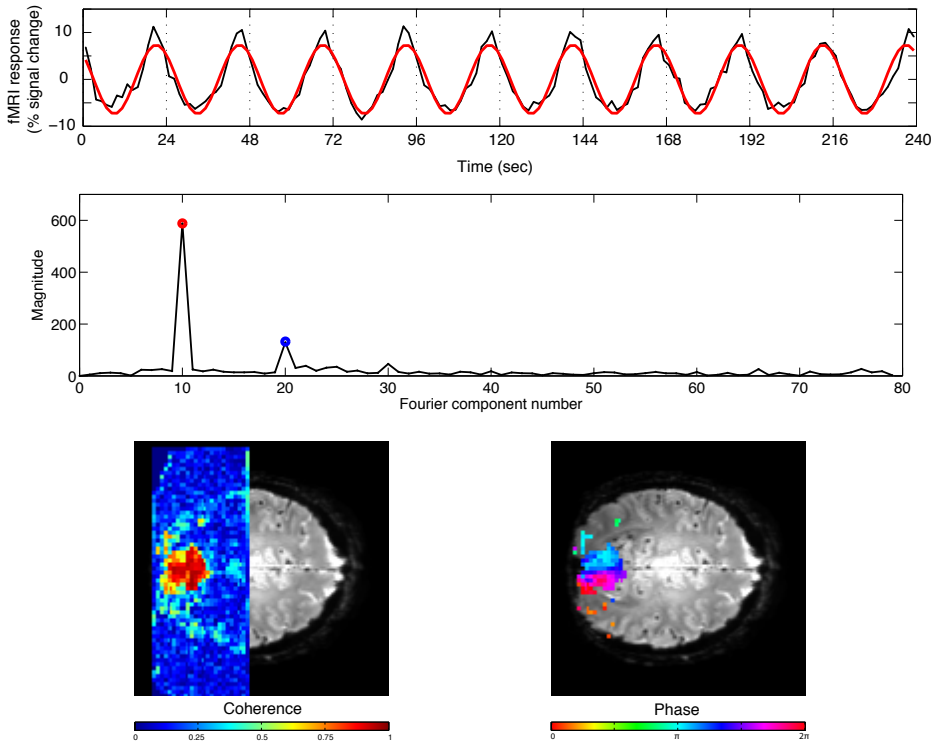

Figure 2.6: Examples of the analysis of retinotopic mapping data. Top Row: Timeseries of a voxel in retinotopic visual cortex in response to the mapping stimulus, and the fitted sinusoid at the stimulus repetition frequency. Middle Row: Fourier decomposition of the timeseries, showing a peak at the stimulus repetition frequency (red circle) and the second harmonic (blue circle). Bottom Row: Coherence (left) and Phase (Right) from the correlation analysis displayed on the anatomy.

of neural activity as the stimulus enters and leaves each receptive field in turn. If the stimulus moves across the visual field at a constant speed, the responses will differ only in their phase. Retinotopic mapping methods that depend on the phase of the temporal modulation to define the receptive fields are often called 'phase-encoded' (Engel et al., 1997).

Figure 2.6 demonstrates how these methods works. The top row shows a timeseries from a voxel in retinotopic cortex in response to the polar angle mapping stimulus, along with the fitted sinusoid at the stimulus repetition frequency. The middle row shows a Fourier decomposition of the timeseries, with a peak at the stimulus repetition frequency, in this case 10 cycles/scan (red circle). There is a second peak at 20 cycles/scan (blue circle), which is a harmonic of the fundamental frequency. To judge the strength of the response to the mapping stimuli, we use the coherence value, which

is the amplitude of the response at the stimulus frequency divided by the total energy over all frequencies. An example of the voxelwise coherence values in response to the polar angle mapping stimulus are shown in the left hand figure on the bottom row. Elevated coherence values are seen in visual cortex. The phase of the sinusoid at the stimulus repetition frequency that best fits a voxels timeseries defines which location in the visual field that voxel best responds to, and can be used to colour code each voxel (Figure 2.6, bottom row). Visual cortices in left and right visual hemisphere can be seen to have different colours because these areas respond to the contralateral visual field.

In both the eccentricity and polar angle mapping conditions, we collected half the scans with the stimuli moving in one direction (e.g. clockwise for the polar angle stimuli, and expanding rings) and half the scans with the opposite direction (counter-clockwise for polar angle mapping, and contracting rings). By combining across these different directions, we could estimate the phase of the response for each voxel independent of any lag in the BOLD signal caused by hemodynamic delay (Larsson and Heeger, 2006).

Eccentricity

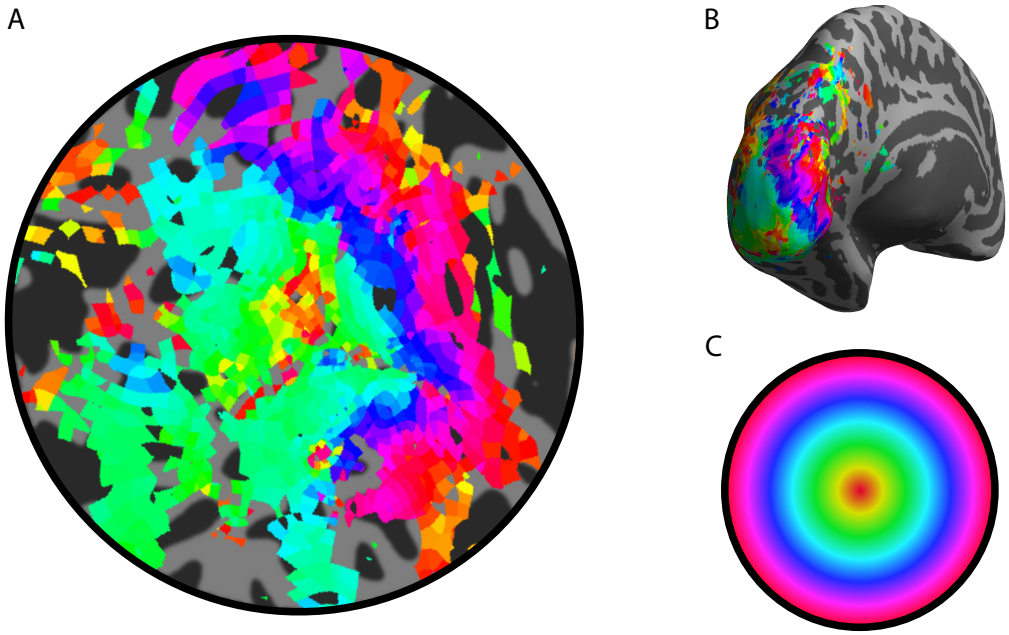

Figure 2.7: Example eccentricity map displayed on a flattened cortical patch for a single subject (A), and on the inflated surface (B). The parts of visual field corresponding to each phase value are shown in panel C.

An example of the eccentricity maps generated by the expanding and contracting ring stimuli is shown in Figure 2.7 on both a flattened patch of cortex (A) and the inflated surface (B). The colour at each cortical location indicates the point in space that area best responds to (C). The fovea is represented in posterior locations, centered at the occipital pole. In more anterior areas, the representation shifts towards the periphery.

Polar Angle

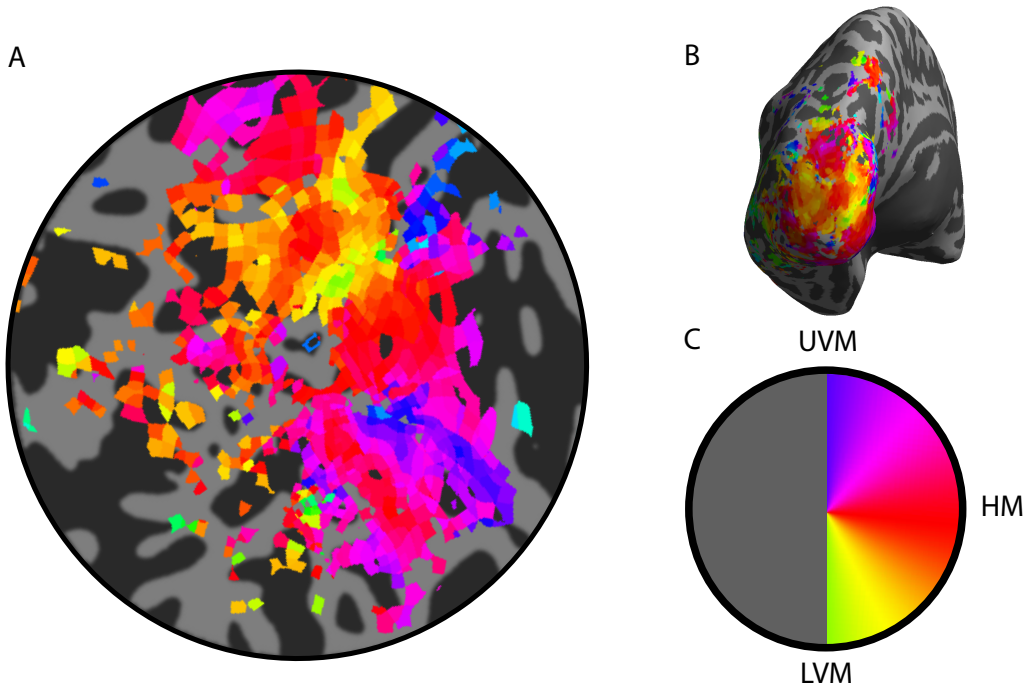

Figure 2.8: Example polar angle map displayed on a flattened cortical patch for a single subject (A), and on the inflated surface (B). The parts of contralateral visual field, from the Upper Visual Meridian (UVM) to the Horizontal Meridian (HM) and Lower Visual Meridian (LVM), corresponding to each phase value are shown in panel C.

The polar angle map generated by the rotating wedge stimuli is shown in Figure 2.8 on a flattened patch (A) and the inflated surface (B). The colour in each area corresponds to the visual angle in the contralateral visual field encoded at each position. Note that the representation of the contralateral visual field is inverted in V1. The horizontal meridian (HM) is represented at the centre of the calcarine sulcus. Moving dorsally towards the upper bank of the calcarine, the representation shift towards the lower vertical meridian (LVM), and then back towards the HM. Moving ventrally towards the lower bank of the calcarine, the representation shifts towards the upper vertical meridian (UVM), and back again. The reversals in direction of the phase mapping mark the borders between the visual areas.

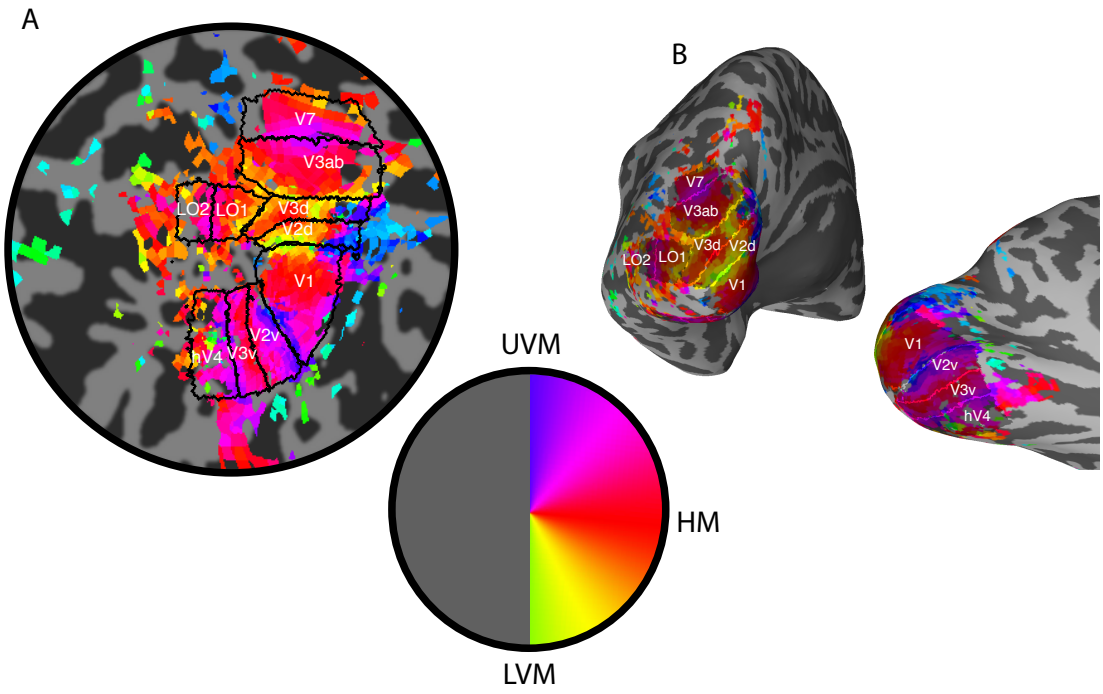

Figure 2.9: Polar angle map plus identified visual areas overlayed on a flattened patch (A) and the inflated surface. Reversals in the polar angle phase map were used to identify the borders between visual areas.

2.2.1 Identifying visual areas

The eccentricity and polar angle maps can be used to identify the borders between the early visual areas (Figure 2.9). V1 contains a full representation of the contralateral visual hemifield, with the contralateral HM represented at the centre of the calcarine sulcus. Area V1 is bordered by two cortical regions that form the dorsal and ventral halves of area V2 (V2d and V2v). Each of these regions represent one quarter of the visual field, V2d representing the lower visual field and V2v representing the upper visual field. The border between V1 and V2 lies at the representations of the LVM and UVM, where the direction of phase progession reverses, and moves back towards the HM. This pattern of reversal continues, with adjacent visual areas having mirrored polar angle maps travelling dorsally and ventrally away from V1. For example the V2/V3 border is marked by a phase reversal at the representation of the HM. The dorsal and ventral V3 maps then progress back towards the LVM and UVM respectively. V1, V2 and V3 share a foveal representation that makes it difficult to differentiate the areas at the most foveal part of the map, known as the area of foveal confluence.

The mirrored phase representation found in V1-V3 changes slightly beyond this point, as non-identical visual areas are found on the dorsal and ventral sides. Area V4 borders

V3v at the UVM, and from here the angular map progresses the whole of the contralateral hemifield, rather than just a quarter-field up to the HM. The exact nature of the V4 angular map is a subject of debate, specifically whether the visual hemifield representation is split dorsally and ventrally, or whether a complete hemifield representation exists within ventral V4, with no dorsal portion of V4 in humans (Tootell and Hadjikhani, 2001). For the purposes of this thesis we adopted the definition of hV4 as a full contralateral hemifield representation beyond V3v (Larsson and Heeger, 2006).

On the dorsal side, V3 borders with 2 visual areas at the LVM. The LVM representation bifurcates around halfway along the border of V3d. The border above this bifurcation marks the border between V3 and V3A. V3A contains a full map of the contralateral hemifield (Wandell et al., 2005) extending dorsally. V3A shares a polar angle map with another area, V3B, with the boundary between the two areas being marked by a shared foveal representation separate from the one at the foveal confluence. As this boundary is not always clear, V3A and V3B are often combined into a single region referred to as V3AB (Larsson and Heeger, 2006). V3AB shares its anterior/dorsal boundary with V7 at the UVM, with V7 also containing a full representation of the contralateral hemifield (Press et al., 2001).

The retinotopic maps beyond V3 were initially poorly defined, but recent studies have shown eccentricity and polar angle maps extending up to and through V5/MT. The LVM representation below the bifurcation in the V3d border marks the boundary between V3d and an area known as LO1 (Amano et al., 2009; Larsson and Heeger, 2006). LO1 also contains a full representation of the contralateral visual field, and borders with LO2 at the UVM. Similarly, the anterior border of LO2 is at the LVM. This border is believed to be the posterior border of MT/V5, which has been shown to contain a map of the contralateral visual field (Huk et al., 2002; Kolster et al., 2010). Amano et al. (2009) showed evidence of a full hemifield representation bordering LO2 at LVM which they called TO1, and an additional map with a border at the UVM that they named TO2, identified as a possible candidate for area MST. In the retinotopy data for our subjects however, LO1, LO2 and the retinotopy in MT+ were not always easy to identify, so LO1 and LO2 were not studied, and MT+ was identified with a within session localizer scan.

MT was identified independently for each scanning session, as a uniform area identified from the separate retinotopy session was not always possible. In some scanning sessions a motion retinotopy stimulus was used as a localizer (Huk et al., 2002). The set-up for this stimulus was identical to the rotating wedge stimuli described above, but in this case the wedge shape was defined by moving versus static dots (Figure 2.10).

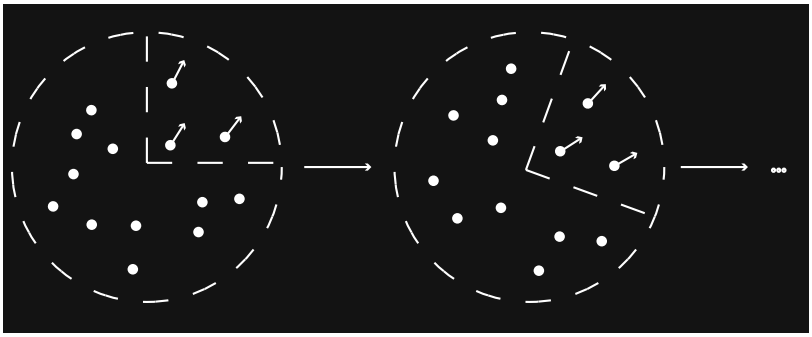

Figure 2.10: The motion defined retinotopy stimulus used as a within session MT+ localizer in some sessions. Other sessions used the same arrangements of dots, but alternated all dots in motion versus all dots stationary.

The wedge was defined in a 24.6° diameter circular aperture filled with white dots on a black background (dot density 5 dots per degree^2 of visual angle, dotsize = 0.1 deg). At a given point in time, dots within a 90° wedge moved toward and away from fixation (dot speed 8 deg/sec), while the rest of the dots within the aperture remained stationary. The wedge advanced 15° every second, completing a full revolution each 24s, and completed 5 revolutions each scan (total, 2 minutes per scan). In several other sessions, an MT+ localizer was used (Huk et al., 2002). This contrasted moving versus static dots for the whole field. This stimulus is known to yield activity in MT, MST and sometimes cells in the LO areas (Kolster et al., 2010).

The use of within session retinotopy scans allowed us to compare retinotopy data collected at different resolutions and field strengths. The initial retinotopy data were collected at 3T at a resolution of 3 mm isotropic, additional sessions at 7T with the motion retinotopy stimulus were collected at 1.5 mm isotropic, and for one subject we collected an additional session of motion retinotopy at 7T with a resolution of 0.8 x 0.8 x 0.85 mm^3. Figure 2.11 shows the retinotopic maps for polar angle collected for the same subject in each different session displayed on the cortical surface. The phase reversals indicating the borders between visual areas are easy to identify at 3 and 1.5 mm, and appear to match up well. The high resolution phase map is less clear than at coarser scales, but the reversals at the edges of the calcarine sulcus can be seen. In both of the 7T datasets, the polar angle map in MT+ is clearer than at 3T, with a full map of the contralateral visual field apparent. The phase reversal between hMT and hMST can possibly be identified, which is not apparent in the 3T dataset. This is an example of

the potential benefits of high-resolution fMRI, although also highlighting the reduced SNR at very high resolutions.

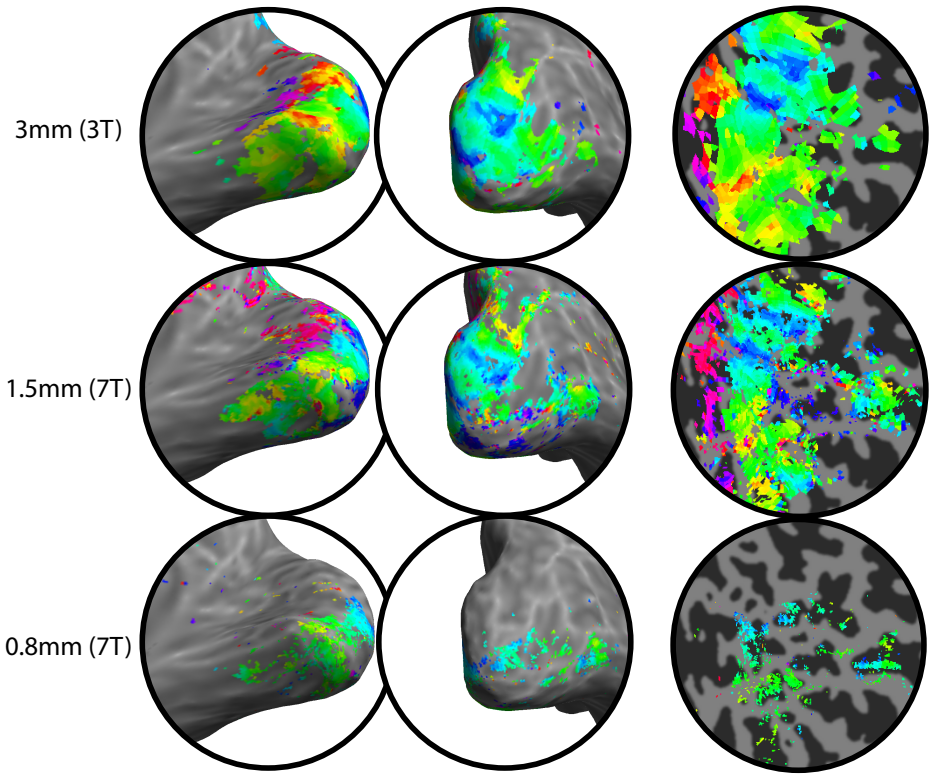

Figure 2.11: Comparing retinotopic maps for polar angle collected in 3 different sessions: a 3 mm isotropic session at 3T, a 1.5 mm isotropic session collected at 7T, and a 0.8 x 0.8 x 0.85 mm^3 session also collected at 7T. Phase values from voxels with a coherence > 0.5 are displayed on the inflated (left) and flattened (right) cortical surface.

2.3 fMRI Methods

2.3.1 Participants

All volunteers took part in the studies with written consent. Procedures were approved by the Medical School Research Ethics Committee of the University of Nottingham.

2.3.2 Visual Stimuli

Stimuli were generated using MATLAB (MathWorks) and MGL
(available at http://justingardner.net/mgl) on a Macintosh Intel computer. Stimuli were displayed via an Epson EMP-8300NL LCD Projector onto a back-projection screen close to the bore of the magnet. Subjects viewed the projected stimuli through a set of mirror-glasses (maximum eccentricity of 16.2 degrees of visual angle).

2.3.3 Functional Imaging

BOLD data were collected on a Phillips 7T scanner, using either a 3D Gradient Echo EPI pulse sequence (TE=25ms, TR=85ms, FA =22 deg dynamic scan time=2s), or a 2D Gradient Echo EPI Sequence (TE=25ms, TR=2000ms, FA=78 deg), both with a voxel size of 1.5 mm isotropic and 34 slices through the brain. Initial piloting was done at 1.25 mm isotropic 3D GE EPI, but it was decided that this size did not have a high enough signal to noise ratio (SNR) to accurately measure BOLD signal. This highlights an important consideration in high-field scanning: the temptation may be to aim for the highest possible resolution to a) increase the localisation of signal and b) cut down on partial volume effects by increasing the likelihood of voxels containing only one kind of tissue, and ideally resolve a single column of neurons with the same selectivity. However, smaller voxels lead to a reduced amount of signal, which leads to a reduced SNR for some experimental paradigms that are not designed to maximise it. After initial higher-resolution imaging failed to resolve the HDR in event-related pilots, the voxel size was therefore increased to 1.5 mm isotropic.

Distortions caused by inhomogeneities of the main (B0) magnetic field are especially problematic at 7T, and without correction the resultant brain images would be unusable. Before the scanning session began proper, it was necessary to take a measure of the B0 magnetic field, and calculate shimming values to combat distortions in the field. Test EPI images with the calculated shim values were manually examined, and retaken with altered parameters if necessary. This process, whilst essential for collecting good

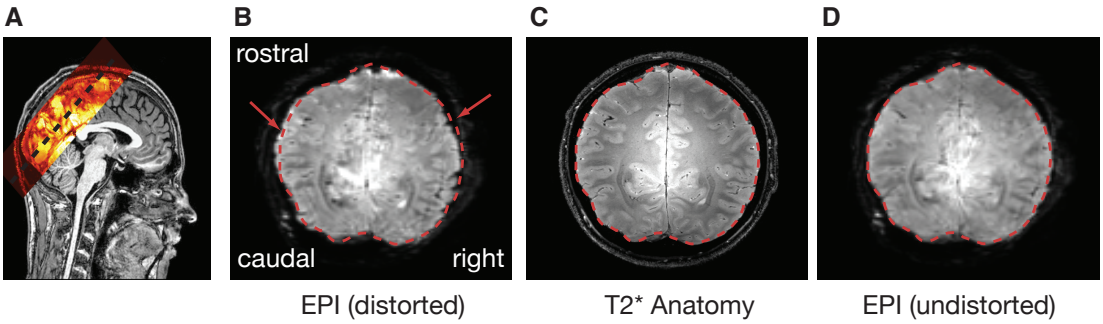

Figure 2.12: Examples of the images collected in a typical scan. A) Example of a whole head MPRAGE image, with typical stack placement in orange. The stack was placed to encompass the calcarine sulcus and temporal and dorsal visual areas B) Example EPI image collected during an experiment. The dashed line shows where the distortions cause the collected image to differ from the 'actual' image. C) T2* weighted anatomical inplane image. This image has the same contrast as the EPI images, but doesn't have the distortions D) EPI image after nonlinear alignment with the T2* weighted anatomy. This has corrected some of the distortions in the image, which would allow the data from the EPI images to be displayed on the whole head anatomy or cortical surfaces.

quality data, can potentially add 10-15 minutes to the scan time, which can be problematic if a large amount of functional data has to be collected. Subjects in the scanner for a prohibitively long session may become restless, leading to increased subject motion and reduced attention for the functional experiments.

As a first analysis step, functional scans were aligned with an inplane anatomy image collected at the end of the scanning session. Initial piloting used a T1-weighted image (MPRAGE, voxel size = 1 x 1 x 1.5 mm^3). In later sessions a high-resolution (0.5 $\times$ 0.5 $\times$ 1.5 mm) T2*-weighted anatomical volume was collected during each scan session. These images had the same slice prescription and coverage as the functional data (TE = 11 ms, TR = 608 ms, FA = 32°). These scans took approximately 2 min. The T2* anatomy was used for nonlinear registration between the distorted EPI images and (intrinsically undistorted) anatomy for each subject, and to calculate the warp fields necessary to correct for distortions (Klein et al., 2009). We used the nonlinear registration as a final step for visualization to project statistical maps calculated in the classification analysis (in the space of the originally acquired data) to be displayed on the subjects' flattened cortical surfaces (in the undistorted anatomical space). Using this correction step also allowed ROIs defined in (undistorted) anatomy-space to be used when analysing the (distorted) 7 T EPI (BOLD) data.

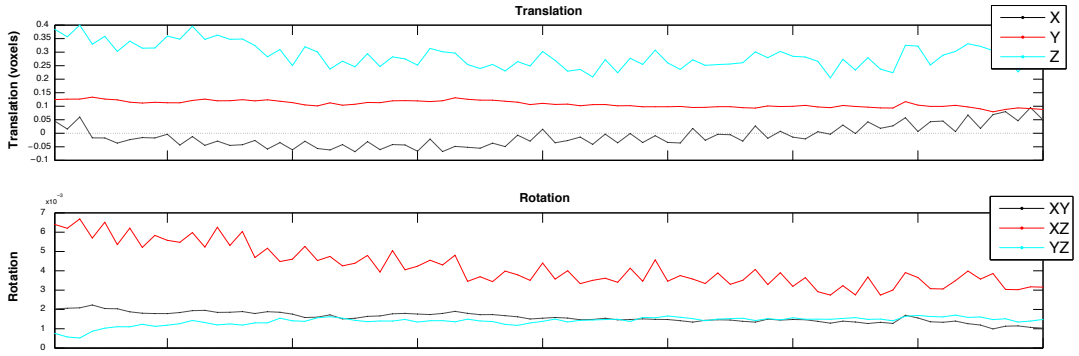

Figure 2.13: Example motion parameters for a single run for a single subject, showing translation (Top Row) and rotation parameters (Bottom Row).

2.3.4 Attention Control Task

During each scan, the subjects performed a demanding two-interval forced-choice contrast discrimination task at fixation to control for changes in attentional state, which are known to modulate fMRI responses (Huk et al., 2001). The attention-control task consisted of a two-interval contrast discrimination task at fixation, with the threshold determined by a staircase. The timing of the intervals and response was separate from the timings of the functional tasks.

2.3.5 Data Analysis

FMRI data were motion-corrected within and between scans for each session using standard motion correction algorithms (Nestares and Heeger, 2000). A rigid-body transform was applied to align each volume to a reference volume, using 6 parameters (3 displacement parameters and 3 rotation parameters). Motion correction parameters for a single run for a single subject are shown in Figure 2.13. Motion correction statistics for each subject in the initial motion classification study are shown in Table 2.1, the mean Euclidean displacement within a single run, and the maximum frame-to-frame Euclidean displacement across a scanning session. Motion parameters for each subject and each session were manually examined to ensure no large motion-shifts that could interfere with successful classification (e.g. shifts larger than the size of a voxel).

Linear trend removal was applied to the to the time course at each voxel. The time-series from the voxels were converted to units of percent signal change by subtracting and dividing by the temporal mean for each voxel across each scan. The scans were then concatenated into a single 4D image. For fast-event related paradigm experiments, additional temporal high-pass filtering was performed. For the block design classifi-

	ab	jb	jp	rs	sh
mean across run (mm):	0.55214	0.92598	0.65225	0.54903	0.76045
max frame to frame (mm):	0.51357	0.49137	0.21273	0.15915	0.44233

Table 2.1: Motion statistics for the initial classification experiment, showing mean displacement across run (Top Row), and max frame to frame displacement across the entire experiment (Bottom Row).

cation experiments, no temporal smoothing or filtering was performed, although individual scans were detrended.

Regions of interest for early visual areas were drawn from the separate session of retinotopy at 3T (described above), to allow identification on the undistorted cortical surface. The ROIs were transformed into the (distorted) EPI space for each subject and scanning session using the nonlinear alignment between the EPI images and the T2* weighted anatomy.

Deconvolution Analysis/Event Related Experiments

Because the potential shape of the hemodynamic response function for each voxel in response to a brief presentation of motion was not known, a deconvolution analysis was used to estimate the response of voxels in visual cortex (Gardner et al., 2005). This method calculates a voxelwise average response to each stimulus type, and assumes that any overlap in responses results in a linear combination (Boynton et al., 1996), an assumption that holds approximately under many circumstances. The BOLD time course is modelled by the following equation:

$$[S_1\, S_2 \ldots S_n][H_1\, H_2 \ldots H_n]^T + noise = [BOLD]^T \tag{2.3.1}$$

where S_i is the ith stimulus convolution matrix with dimensions M x N, where M is the length of BOLD timeseries in TRs, and N is the length of the HDR we wish to calculate in TRs. A deconvolution matrix is a Toeplitz matrix with a 1 in the first column for any TR in which that stimulus was present. For example, a stimulus convolution matrix to calculate a 5 TR HDR for a stimulus presented on the 1st and 7th TR would take the form:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ & & ... & & \end{pmatrix}$$

H_i is the 1 x N unknown HDR to the ith stimulus, T is the transpose operation and any noise is assumed to have gaussian distribution with zero-mean. BOLD is a 1 x M array containing the demeaned BOLD timeseries for a given voxel. The H_i for each stimulus is then computed for each voxel that minimizes the squared error between each side of the equation. An estimated timecourse could then be calculated for each voxel by multiplying the stimulus convolution matrices (S) with the estimated HDRs (H), and the amount of variance in the original timeseries explained by the model could be computed using the equation:

$$r^2 = 1 - \frac{variance(residual)}{variance(original)} \tag{2.3.2}$$

where the residual timeseries is the difference between the estimated and original timeseries. This generated a voxelwise map of r^2 indicating areas of the brain whose activity was modulated by the stimulus. The statistical significance of a given r^2 was calculated using a permutation procedure (Gardner et al., 2005). The deconvolution analysis was repeated with randomised stimulus onset times, so the the times were no longer time-locked to measured activity. The r^2 calculated from the HDRs found using these randomised times were assumed to represent the null distribution of r^2 expected by chance correlations of noise in the timeseries with the stimulus timings. Any r^2 from the distribution of actual r^2 greater than a given threshold calculated from the null distribution (i.e. the 95th percentile) were considered to be significant.

Multivariate Pattern Analysis

Multivariate pattern analysis (MVPA) was used to classify the patterns of activity in each ROI according to motion direction. The input patterns were created by averaging the activity in each voxel across each of the stimulus blocks, leading to one value per direction per run for each voxel. Voxels from each ROI were chosen on the strength of their response to the within-session (but independent) localiser stimulus, by selecting the 500 voxels with the highest coherence values.

Support Vector Machine

We utilized the LIBSVM toolbox to classify the observed directions of motion (Chang and Lin, 2011). A 'linear ensemble detector' for direction of motion was used to calculate a weighted sum of the voxel intensities in the input patterns: this was done by optimizing the weights such that the detector for each direction maximized its output for its preferred direction, and reduced it for other directions. The calculation for a linear ensemble detector for direction θ_k is a linear function of voxel inputs $\mathbf{x} = (x_1, x_2, ..., x_d)$

$$g_{\theta_k}(\mathbf{x}) = \sum_{i=1}^{d} w_i x_i + w_0 \tag{2.3.3}$$

where w_i is the weighting for each voxel i and w_0 is the bias. To construct the required weighting function for each direction, linear discriminant functions were calculated for each pairwise combination of directions using linear support vector machines (SVM). These pairwise discriminant functions were a weighted sum of the voxel inputs in the patterns, satisfying

$$g_{\theta_k\theta_l}(\mathbf{x}) > 0 \tag{2.3.4}$$

$$g_{\theta_k\theta_l}(\mathbf{x}) < 0 \tag{2.3.5}$$

g being greater than 0 if $\mathbf{x}$ is a pattern of fMRI response induced by direction θ_k, and less than 0 for direction θ_l. A linear SVM found the optimal weights and bias for each discriminant function based on the training data set. The pairwise functions for a given direction versus all other directions were added together to yield the linear detector function for that direction.

$$g_{\theta_k}(\mathbf{x}) = \sum_{m \neq k} g_{\theta_k\theta_m}(\mathbf{x}) \tag{2.3.6}$$

This function becomes greater than zero when the input vector $\mathbf{x}$ (voxel pattern of fMRI response) is one induced by its preferred direction θ_k. The *test* patterns are assigned to a given direction label based on which detector function yielded the largest output.

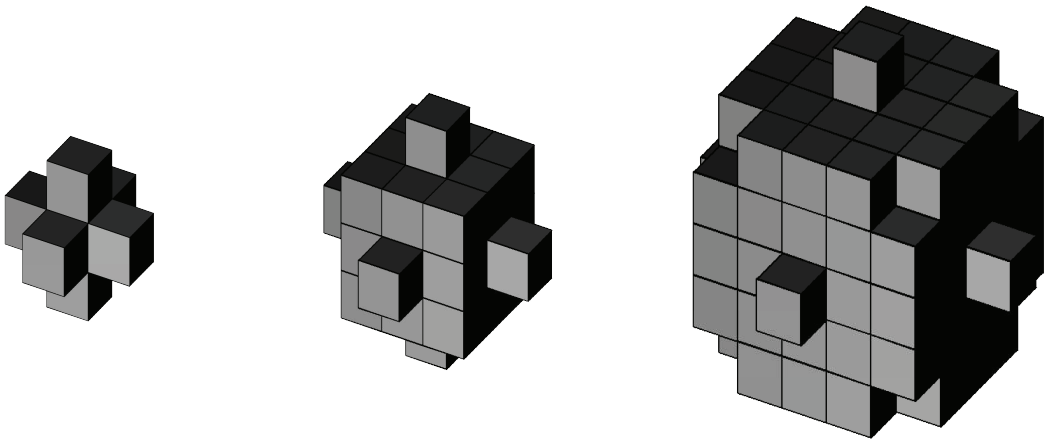

Figure 2.14: Examples of the 3D 'spherical' searchlights. Using a searchlight with small radius (left) increase the precision with which informative areas can be located, but runs the risk of including too few voxels for successful classification. Increasing the radius of the searchlight (centre and right) increases the number of voxels available for classification, but reduces the precision with which informative areas of the brain can be localized.

Cross Validation To test the accuracy of the classifier, we used 'leave one run out' cross validation. The classifier was trained using the data on all but one of the runs, and then its accuracy at classifying the patterns from the remaining run was assessed. This process was repeated for each run in turn. This process assures the statistical independence of the test and training data, as they are drawn from separate runs.

Significance Testing To test whether the classification accuracy in each ROI analysis was significantly above chance, we performed a non-parametric permutation test. 1000 classification analyses (using leave one-run out cross validation) were performed with the labels in the training data indicating which patterns of activity related to which direction of motion shuffled. This produced a distribution of classification accuracies expected under the null hypothesis that these patterns did not relate to the perceived direction of motion. These null distributions were used to create significance values for the actual accuracies, and accuracies from the correctly labelled dataset were considered significant if they were higher than the 95th percentile of this null distribution ($p < 0.05$, one tailed permutation test).

Searchlight Methods A method of generating patterns for classification that does not rely on ROI identification is the searchlight method (Kriegeskorte et al., 2006). In this method a restricted searchlight (normally a sphere in terms of the 3D anatomy) is centered on a given voxel and all voxels that fall within the searchlight are used to create the patterns for classification (Figure 2.14). Standard cross-validation is then used to

generate a classification accuracy, which is assigned to the voxel at the centre of the searchlight. The searchlight is then centered on a different voxel, and the classification analysis is repeated. By applying the searchlight to each voxel in the brain in turn, a map of accuracies can be generated, indicating areas containing information about the stimulus. The size of the searchlight can be altered (i.e. for a spherical searchlight the radius can be increased or decreased) to find the balance between including enough voxels for successful classification and using a small enough searchlight that information can be accurately localized. Figure 2.15 shows the output from a classification analysis using searchlights of varying radius displayed on the inflated cortical surface. Using a searchlight containing a single voxel shows no elevated accuracies in any visual area. As searchlight size increases, areas of visual cortex begin to show elevated accuracies, with accuracy increasing as the searchlight size is increased further. The maps of classification accuracy can then be tested for significance, for example with a binomial test for above chance classification that is then corrected for multiple comparisons (e.g. false discovery rate). We used the searchlight method in a number of experiments to identify areas with high information content to then study them in more detail with ROI based methods.

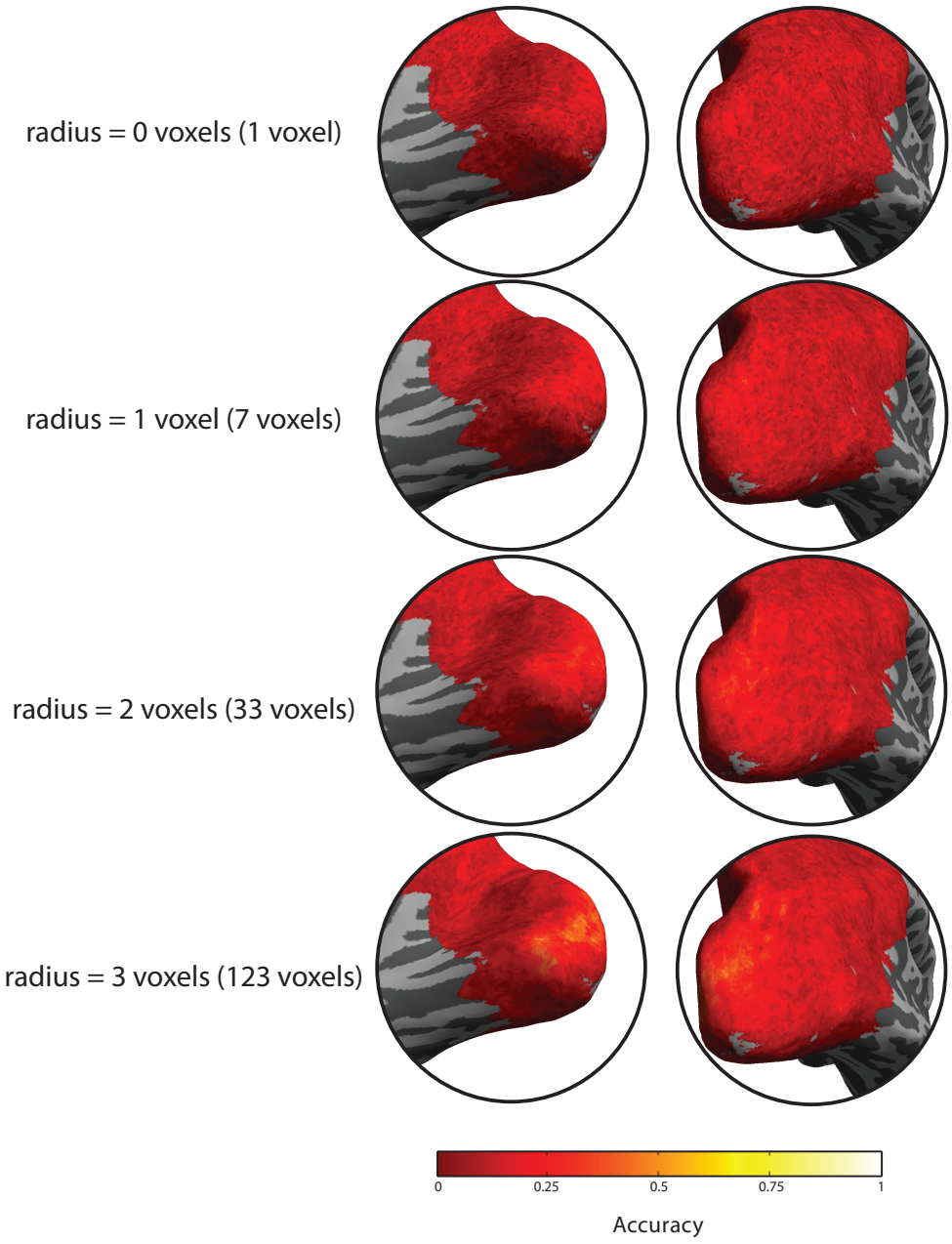

Figure 2.15: Classification accuracy maps of perceived direction of drifting dot patterns, showing proportion of stimuli classified correctly when a searchlight is centered at a voxel on that location on the cortical surface (8 stimulus classes, chance = 0.125), derived from a searchlight classification analysis using spherical searchlights of varying radius. As searchlight size increases, elevated classification accuracies are found.

Chapter 3

Pilot Experiments: Direction Selectivity in MT+

Area MT in the macaque has been called the 'motion area', due to its high proportion of direction-selective cells (Albright et al., 1984; Dubner and Zeki, 1971; Zeki, 1980). An area in human cortex known as MT+ has been suggested as the human homologue of this 'motion area', based on the effects of brain lesions around this area of the brain (Zihl et al., 1983). Non-invasive neuroimaging techniques such as fMRI have also indicated a preference for motion versus static stimuli in this area (Watson et al., 1993).

One major difficulty in studying neural selectivity is the scale on which fMRI records responses, combined with the fact that fMRI involves an indirect measure of neural activity (the BOLD signal). Voxel dimensions in fMRI experiments range from around 1.5 to 3 mm, and these voxel sizes contains tens of thousands of individual neurons. Electrophysiological work has shown that neurons with similar selective properties (i.e. for the same direction of motion) are arranged in close proximity in 'columns' (Albright et al., 1984), but even these neural features are generally too small to be resolved by standard resolution fMRI, with an average of ~500 μm for motion direction columns in MT (Born and Bradley, 2005). Therefore in most cases the BOLD signal recorded from a single voxel will reflect the output of a mixture of 'columns' with a variety of different selectivities.

We undertook a series of experiments to investigate which methods were appropriate to study neural selectivity. We first measured BOLD responses to stimuli of varying strength to see if the signal measured at a coarse spatial scale could be related to single neuron activity. We then utilised an fMRI adaptation paradigm to see if stimuli travelling in different directions could be reliably differentiated, and whether this was

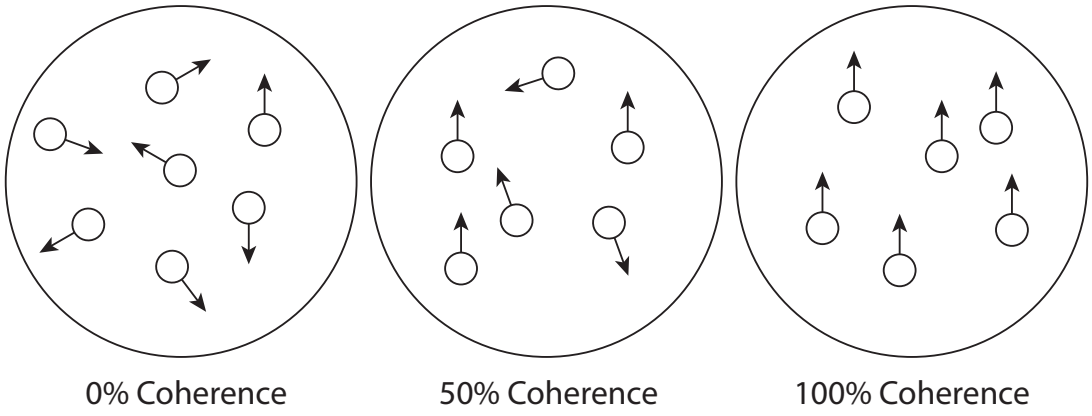

Figure 3.1: Examples of random dot kinematograms (RDKs) with increasing levels of motion coherence.

dependent on the angular separation between the directions of motion.

3.1 Coherence Response Curves

Many neurons that are selective for a particular stimulus property modulate their activity on the strength of the perceptual signal as well as its closeness to its preference. For example, neurons in macaque V1 increase their activity with increasing contrast for sinusoidal grating stimuli that matches their preferred orientation (Movshon and Newsome, 1996). Comparisons of an increase in BOLD signal in human V1 and neural firing rate in macaque V1 with increasing contrast found the rates of increase were comparable (Heeger et al., 2000).

Motion-selective cells in macaque MT have an analogous sensitivity to motion coherence in a random dot kinematogram (RDK) (Figure 3.1). At 0% coherence (all dots in a random dot display moving randomly), all direction selective cells will fire equally, as all directions of motion are equally represented. As motion coherence increases (the number of dots moving in the same direction increases), cells with that preferred direction will increase their firing rate linearly (Britten et al., 1993). Cells which have a preferred direction opposite to the coherent motion will reduce their firing rate, whilst the firing rates of cells with other preferred directions will remain roughly unchanged. The number of cells that increase their activity for their preferred direction is generally greater than the number of cells with the opposite preference that suppress it, leading to a net increase in average neuronal response. The activity of these direction selective cells is closely correlated with the psychophysical performance of the animals (Britten et al., 1992), lesions of macaque MT will raise motion coherence thresholds (the

amount of motion coherence required for successful motion discrimination) (Newsome and Pare, 1988), and stimulation of cells in this area can bias the perceived direction of RDKs in a discrimination task (Salzman et al., 1992). Together, these results indicate a role for MT direction selective neurons in the mediation of perceived motion.

Evidence from lesion studies (Zihl et al., 1983) and fMRI studies contrasting static and moving stimuli (Watson et al., 1993) indicate that MT+ is the human homologue of this macaque area, and should therefore show the same response properties for coherent stimuli. Some studies have shown a preference for coherent over incoherent motion stimuli in MT+ (Braddick et al., 2001; Rees et al., 2000), similar to the increase in mean firing rate found in macaque MT neurons for coherent versus incoherent motion. Some additional areas showed a preference for incoherent motion, whereas area V1 in fact showed a preference for incoherent motion (Braddick et al., 2001). In addition, Rees et al. (2000) were able to show a steady (linear) increase in BOLD signal as the coherence of the dot stimuli was increased, providing further evidence linking BOLD signal to the activity of direction selective neurons. Other studies however have been unable to find such a preference in MT+ for coherent versus incoherent motion (McKeefry et al., 1997; Smith et al., 2006), or have suggested that the different subregions of the MT+ complex may have different coherence sensitivities (Becker et al., 2008).

As discussed in Section 1.3, many neurons in the visual system will change their response properties after repeated or extended presentation of a given stimulus (the 'adapter'), with consequences for perception of subsequent stimuli (the 'probe'). The most well known adaptation effect for motion is the motion after-effect (MAE), where extended viewing of a moving stimulus will lead to a perception of motion in the opposite direction when a stationary, or incoherent motion stimulus is subsequently viewed. The effect is believed to result from a reduction in the activity (the so-called 'fatiguing') of neurons selective to the direction of adaptation, which leads to an imbalance in the sensitivity of neurons tuned to different directions. Elevated BOLD activity in area MT+ coinciding with the MAE was presented as evidence this was the locus of the effect (He et al., 1998), and hence as evidence for direction selective neurons in human MT. However, subsequent studies that controlled for the attention grabbing properties of the MAE found a reduced (but still present) effect (Huk et al., 2001). Activity in area MT measured with fMRI is known to be modulated by attention, so it is important that any study of this area controls for attention, such as by using the contrast discrimination task at fixation in our study. The neuronal responses of direction-selective cells in monkey MT have been shown to be reduced following adaptation with motion stimuli in their preferred direction (Van Wezel and Britten, 2002), which changes the shape

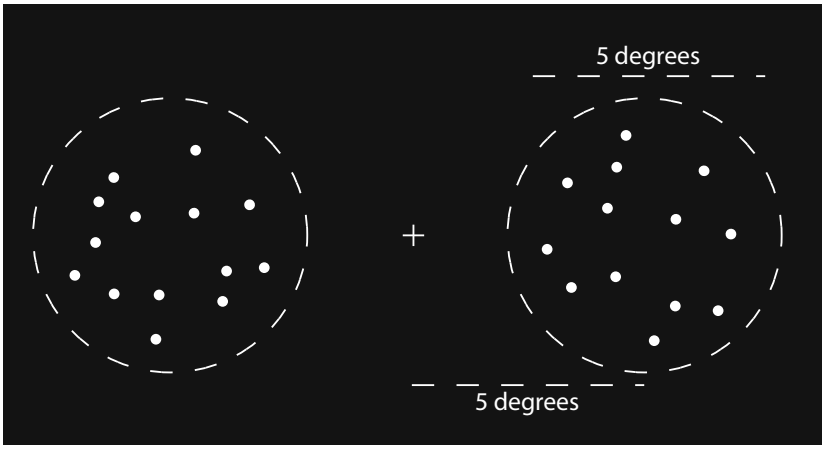

Figure 3.2: Example of the dot stimuli used in the FMRI pilot experiments.

of the coherence response curves measured from these cells. Huk et al. (2001) found reduced BOLD responses in MT+ (and other areas) to a probe stimulus that matched the direction of an adapting stimulus, which they took as arising from a reduction in response level by direction-selective neurons in the brain.

As part of our initial piloting, we tried to measure a 'coherence response curve' from voxels in visual cortex under both 'unadapted' (incoherent motion in between trials) and 'adapted' (coherent motion in between trials) to see if we could find the ideal adapter/probe coherences to use in a later adaptation experiment. We recorded BOLD responses to dot stimuli moving at a range of coherences in an upward direction, from 0% (incoherent motion) to 100%.

3.1.1 Methods

Stimuli for all experiments consisted of two patches of dots (patch diameter = 5 degrees) centred 5 degrees from fixation in the left and right visual fields (white dots on a black background, dot density = 5 dots/deg^2, dot speed = 8 deg/sec, dot size = 0.1°) (Figure 3.2). This stimulus set-up followed that used by Hol and Treue (2001) in psychophysical studies of motion adaptation, and kept the dot patches away from fixation so individual dot tracking could not be used by subjects. During the event related scans, these dots would move in various directions in bursts of 1 second with varying levels of coherence (the percentage of dots moving in a single common direction) or direction, with an ISI of variable length where the dots either moved randomly or remained static. The direction and coherence level of each dot patch could be controlled

individually if necessary. Separate dot patches were used instead of a single central dot patch to allow a separate task at fixation to control for attention, and allow separate adaptation of left and right visual field. For the dot coherence experiment, probe stimuli consisted of upwardly drifting dots with coherence of 25%, 50%, 75%, 100% or 0% (incoherent motion). For half the scans, incoherent dot motion was displayed in both patches during the adaptation periods. For the remaining scans, bidirectional motion (50% dots drifting upwards, 50% drifting downwards) was shown in the right dot patch during the adaptation periods, whilst incoherent motion was shown in the left dot patch. A single subject also undertook a second session of scanning in which static dots were displayed during the adaptation periods.

Data Analysis

Data were aligned and motion corrected as described in Chapter 2. In addition, data were high-pass filtered, with a cut-off frequency of 0.01 Hz to suppress slow signal drifts, whilst preserving the high frequency signal in response to stimulus presentation. Signal was converted to percent signal modulation by dividing by the temporal mean for each voxel.

Because the potential shape of the hemodynamic response function for each voxel in response to a brief presentation of motion at each coherence level was was not known, a deconvolution analysis was used to estimate the response of voxels in visual cortex (Chapter 2, Gardner et al. (2005)).

ROIs were created on the basis of responses to a retinotopic localizer. A contiguous group of voxels with a coherence > 0.4 in the lateral part of the brain were considered to correspond to the left and right portions of MT+. The voxels used to calculate the HDR to the probe stimuli were selected on the basis of the r^2 thresholding method described above. For the initial coherence curves experiment, a p of 0.05 was used as the threshold for voxel selection.

For the 'unadapted' scans, voxels from right and left MT+ were combined. For the adapted scans, left and right MT+ were examined separately.

3.1.2 Results & Discussion

Figure 3.3 shows the calculated hemodynamic response curves from voxels in MT+ for a single subject for scans where the subject viewed incoherent dots during the adaptation periods in between trials. The curves show the expected pattern with the 100%

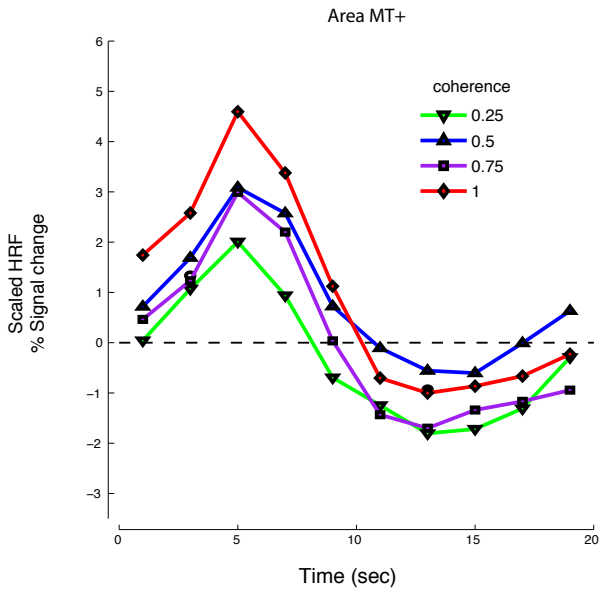

Figure 3.3: HDR curves in MT+ for different coherence levels for a single subject using a deconvolution analysis

coherence curve having a largest amplitude, and the 25% response being the lowest, indicating a preference for coherent motion in MT+ as in previous studies (Becker et al., 2008; Braddick et al., 2001; Rees et al., 2000). The intermediate coherences lie in between the two curves, although the are not fully separable or ordered, with the 50% and 75% curves overlapping. This may just be due to variability in the responses to these stimuli, or it may be that the balance between the signal caused by activity relating the preferred, anti-preferred and other directions may not allow these signals to be differentiated (Braddick et al., 2001).

Figure 3.4 shows the mean HDRs from right (incoherent motion adapted) and left (bidirectional motion adapted) MT+ ROIs in 2 subjects during the adaptation condition. The pattern suggested by the curves from the single subject remains in the 'incoherent motion adapted' ROI, with the amplitude for the 100% coherence curve larger than for those of lower coherences. The curve for 0% coherence is lower still, although a slight increase in BOLD signal remains. This may be due to the fact that whilst a 0% coherent stimulus should not lead to any response in the 'incoherent motion adapted' ROI (as it matches exactly the adapting stimulus), it will lead to a response in the 'bidirectional motion adapted' ROI, which may lead to a general increase in attention for all trials. As changes in attentional state are known to affect activity in MT+ (Huk et al., 2001), this is a possible explanation for this anomalous activity. Nevertheless, the activity does not match that for the higher coherence stimuli. The curves for the inter-

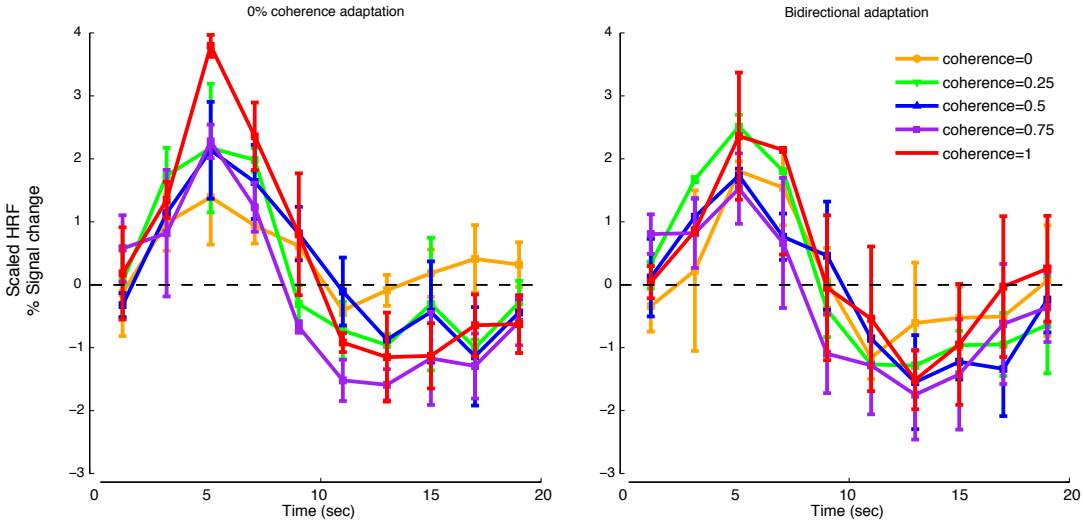

Figure 3.4: HDR curves for different coherence levels using a deconvolution analysis from MT+ during the adaptation condition for the 'incoherent motion adapted' and 'bidirectional motion adapted' MT+ ROIS. Errorbars indicate ± 1 SEM across subjects (n=2).

mediate coherences overlap, so no exact ordering of responses can be shown. For the 'bidirectional motion adapted' ROI, the ordering appears to be altered, with responses for low coherence stimuli elevated compared to high-coherence stimuli. However, the large standard error bars make it difficult to draw firm conclusions about the ordering of the coherence curves.

Figure 3.5 demonstrates the potential problems in measuring coherence response curves using fMRI, namely the choice of baseline in between trials. HDR curves for different coherences were measured during a session where a subject viewed stationary dots in between probe stimuli, rather than incoherently moving ones. The amplitude of the responses to coherent, incoherent and intermediate coherence stimuli are all very similar, with some suggestion of larger responses for lower coherence stimuli. Voxels in MT+ responded just as strongly to incoherent motion as they did to fully coherent motion, seeming to indicate no kind of selectivity. However, this highlights the issues in attempting to create 'coherence response curves' related to 'contrast response curves': motion selective cells will still respond to incoherent motion, just less than to coherent motion at their preferred direction. In addition, motion onset can induce a large transient response that can mask any truly selective signal. Therefore, it may be the case that responses to different levels of coherence must be measured in terms of response to a change in coherence (e.g. a chance from incoherent to coherent motion), not by the sudden appearance of motion in a stimulus.

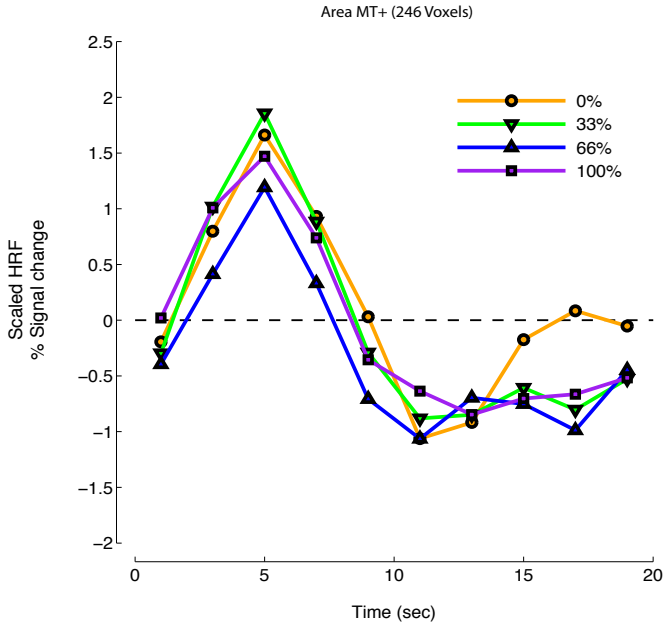

Figure 3.5: HDR curves for different coherence levels using a deconvolution analysis from MT for a single subject using static dots as the adaptation stimulus.

FMRI voxels, even the relatively small ones used in these initial experiments, will sample a large number of motion sensitive neurons, with a range of preferred directions, speeds and other stimulus properties. Although increasing the coherence of a stimulus may increase the activity of cells selective for the coherent direction, cells selective for other directions will reduce their activity as the number of dots moving in their preferred direction decreases, and may even be suppressed by the cells selective for the coherent motion. Combined with the uncertainty about the ratio between a selective signal and a non-selective 'global' component in the BOLD signal, it is not simply the case that an increase in motion coherence will lead to an increase in BOLD signal. For the study of selectivity with neural architecture arranged at a scale less than that of typical fMRI voxels, other techniques may be required.

3.2 Adaptation

Extended exposure to a perceptual stimulus can affect the perception of a subsequent stimulus in a way that can provide evidence regarding the neural underpinnings of the perception of those stimuli (Clifford, 2002). Similarly, the amplitude of the BOLD response to a probe stimulus after extended exposure to an adapting stimulus can tell us about the selectivities of the neurons underlying that BOLD response in a given

voxel or area (Grill-Spector and Malach, 2001).

Outside of perceptual effects, adaptation can also affect the detection and discrimination of subsequently viewed stimuli (Clifford, 2002). Direction and orientation studies have shown that adaptation to a stimulus will make detection of that stimulus more difficult, for example bidirectional adaptation along an axis of motion causes coherence thresholds along that axis of motion to be elevated, even in the absence of an MAE (Hol and Treue, 2001). Adaptation orthogonal to the test direction has no effect, with adaptation increasing as adapter and test become more similar, revealing a direction specific elevation of detection thresholds. For two-interval discrimination the pattern is reversed, with directions close to the adapted direction being easily discriminated, whilst if adapter and test directions differed by around 20 degrees, discrimination thresholds were elevated (Hol and Treue, 2001).

Neural selectivity for *orientation* has been demonstrated in a number of fMRI adaptation studies (Fang et al., 2005; Larsson et al., 2006; Sapountzis et al., 2010). BOLD responses in primary visual cortex to oriented grating stimuli were measured after adaptation to matching, orthogonal and varyingly separated stimuli (Fang et al., 2005). The amount of adaptation was orientation selective, specifically inversely proportional to the angular separation between test and adapter. Psychophysical adaptation tests done alongside the experiment found a similar relationship for contrast detection thresholds after adaptation. It is worth noting that initial studies found fMRI adaptation for orientation only in extrastriate cortex, and not in V1 (Boynton and Finney, 2003). These first studies used a very brief adaptation period (~1s), and when later studies extended this adaptation period to match with psychophysical studies, they actually found more adaptation in primary visual cortex compared to extrastriate areas. Therefore failure to find an adaptation effect does not mean that neural selectivity does not exist in the area under examination, it may be that the stimuli chosen as the adapter may not reveal it.

FMRI adaptation has been used to show that BOLD responses are reduced when motion stimuli follow an adapter with the same direction of motion compared with an opposite one (30 s initial adaptation, 4 s top-up adaptation) indicating some level of direction selectivity (Huk et al., 2001). An additional experiment utilising an alternative adaptation method compared BOLD signal during blocks of stimulus presentation where stimuli moved repeatedly in the same direction with blocks in which a variety of stimulus directions were shown (Huk et al., 2001). Given that psychophysical adaptation has also shown direction-tuned effects for motion stimuli, we decided to investigate neural selectivity for motion direction using fMRI adaptation in a manner similar to the orientation studies by examining whether the amount of adaptation changed

with the angular separation between adapter and probe.

3.2.1 Methods

A 'long-term' adaptation paradigm was used, with a long initial adaptation and a period of 'top-up' adaptation in between trials, as this had yielded good results in the Hol and Treue (2001) psychophysical adaptation experiments for motion and in earlier fMRI adaptation studies. Data were collected from 2 subjects on the 7T scanner, with the same scan parameters and data pre-processing as the pilot coherence study. ROIs were again identified using the retinotopy stimulus used in the pilot experiment, and the same fixation task was used to control attention. The main part of the experiment was an event-related adaptation paradigm. Stimuli consisted of the same dot patches in the left and right visual fields as used in the coherence study. Before the scan began, there was a 30s adaptation period bidirectional motion stimuli (50% moving upwards and 50% moving downwards) on the side to be adapted, and static dots on the unadapted side. Bidirectional motion was used to match the earlier psychophysical studies, and to avoid inducing an MAE that could potentially introduce an attentional confound (Huk et al., 2001). The event related trials during the scan consisted of a period of 'top-up' adaptation (7-13s) followed a 1s motion probe stimulus either matching the adapted direction, or differing by an angular separation of up to 90 degrees (orthogonal), oriented towards fixation. Based on the previous psychophysical and fMRI adaptation experiments, we expected complete recovery from adaptation when probe and adapter were orthogonal to each other, but as the angular separation between adapter and probe is reduced, the amount of shared neural population processing the stimuli should increase, and therefore the reduction in signal due to adaptation should increase.

3.2.2 Data Analysis

The same deconvolution analysis (Chapter 2) used for the coherence experiment was used to generate HDR curves for the different directions.

3.2.3 Results & Discussion

Preliminary results (Figure 3.6) showed some adaptation effects, but it was difficult to establish statistical significance (beyond the trends seen in the data). Although it does appear that the hemodynamic response amplitude for the test direction matching

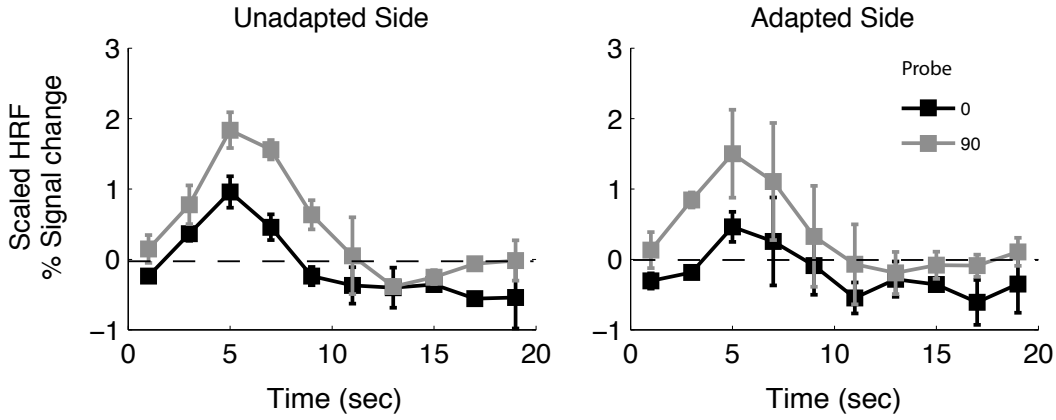

Figure 3.6: Hemodynamic response curves from area MT from the 'unadapted' side (i.e. ipsilateral to the adaptation stimuli) and the 'adapted' (contralateral) side, for the probe directions matching (black line) and orthogonal (grey line) to the adapter stimuli. Errorbars indicate ± 1 SEM (n=2).

adaptation is reduced (dir=0, black curve), with little adaptation seen for the orthogonal direction (direction=0, grey curve), there is a lack of a within session unadapted baseline for comparison. Comparing the adapted hemodynamic response curves with those from the 'unadapted' side shows that the hemodynamic response curves are not equal on the unadapted side, highlighting that we cannot assume in this case that the difference in response levels between the directions are a result of the adaptation. It may be that the fixation task was not fully controlling attention, and some directions modulate attention, or that the adaptation in either visual field was not as spatially specific as we had assumed.

A number of previous studies have demonstrated adaptation effects for motion. Reduced BOLD amplitude was found in MT+ (as well as V1 and V2) when the direction of a drifting grating matched that of an adapting grating compared to a probe grating of the opposite direction (Huk et al., 2001). Contrasting the BOLD signal during blocks of mixed directions of motion and blocks of constant direction showed a reduction in BOLD amplitude during the constant directions block for both gratings and plaid stimuli (Huk and Heeger, 2002; Huk et al., 2001), with the strongest effects seen in MT+. A recent study that used RDK stimuli to probe motion adaptation also found direction-specific adaptation effects in a variety of areas (Lee and Lee, 2012).

The adaptation experiment demonstrates some of the difficulties of studying neural selectivity with fMRI. Previous studies have found varying levels of adaptation based on the stimulus and experimental parameters selected (Fang et al., 2005), with the relationship between the adapter and probe stimuli seeming to be a key concern. Re-

sponse curves recorded from individual cells (stimulus strength on the horizontal axis, strength of response on the vertical axis) show an s-shaped curve, with the firing rate saturating as the stimulus increases in strength (i.e. contrast or coherence). After adaptation, this curve is 'shifted' along the horizontal axis, meaning that a stronger stimulus is required for the cell at the rate it did pre-adaptation. The point at which the adaptation effect will be felt most strongly is on sloped part of the curve, i.e. before the firing rate saturates, meaning that a probe stimulus should be selected that lies on this part of the curve. We elected to use a low coherence (33%) probe, mirroring the low-contrast probes used in orientation experiments (Larsson et al., 2006; Sapountzis et al., 2010), as this strength of stimulus should lie on the desired part of the curve, although the low coherence may have led to a sub-optimal BOLD response. We initially attempted to find the ideal probe coherence in the experiments described in Section 3.1, but were unable to find the ideal coherence from these results. Therefore we tried to use a motion equivalent to the contrast stimuli used in earlier fMRI adaptation experiments, which may not have been the optimal stimulus in this case.

There are also concerns that the measured effects of fMRI adaptation may not measure neural adaptation as understood in electrophysiological experiments, and may in fact indicate selectivity where there is none by shifting adaptation effects 'downstream' (Bartels et al., 2008). Areas without selectivity for a stimulus condition can appear to have it in fMRI adaptation studies due to 'inheriting' it from earlier areas. For example, area V4 demonstrated selective adaptation to motion direction in some studies (Huk et al., 2001), even though this area is normally thought to insensitive to motion. One hypothesis for this result is that while neurons in V4 do not adapt to motion, neurons in V1 (which output to V4) do adapt, meaning that V4 may show a reduction in BOLD response even though there is no adaptation or selectivity in that area. The study by Lee and Lee (2012) highlighted this issue in their study of direction-selective adaptation by showing that the adaptation effects they found in MT and MST could potentially be largely explained by the adaptation effects seen in earlier areas, indicating that the adaptation effects seen in these higher areas may just have been inherited from lower areas, with the possibility of no (or little) additional adaptation in these areas. Evidence for V1 being the source of adaptation effects also comes from the fact that adaptation in MT neurons has been shown to be position specific within the cells RF, to an area within the RF of a V1 input to the cell (Kohn and Movshon, 2003). MT neurons have however been shown to alter their tuning for direction of motion in a fashion not seen in V1 cells, and whether these additional changes reflect additional processing in higher areas or can be explained by inherited effects remains to be seen (Kohn and Movshon, 2004).

The indirect neurovascular coupling between neural activity and the BOLD signal creates additional issues when interpreting adaptation results. The extent to which BOLD activity and neural spiking are linked is unclear, and it has been argued that an increase may reflect inhibitory or excitatory synaptic activity, rather than action potentials (Logothetis et al., 2001). This may lead to areas receiving adapted input displaying selective fMRI adaptation effects rather than the truly selective areas (Bartels et al., 2008), going beyond the 'inheritance' issues discussed above and leading to the apparent locus of neural adaptation being 'shifted' downstream. Adaptation in an fMRI experiment may not be enough to conclusively demonstrate neural selectivity in specific brain areas, but if a consistent adaptation effect can be shown it demonstrates that some kind of selectivity must exist in some part of the visual system

3.3 General Discussion

Summaries of some of the pilot experiments we performed indicate some of the difficulties encountered when studying neural selectivity for motion using standard fMRI methods, even using smaller voxels at a higher field-strength. Although direction-selective cells in macaque MT are known to increase their activity as motion coherence in their preferred directions increase, the change in BOLD response for a given voxel given a change in coherence may not not be straightforward (Becker et al., 2008; Rees et al., 2000; Smith et al., 2006). Incoherent motion contains a large number of motion signals that will weakly activate a large number of direction selective cells. Therefore, although increasing the coherence will cause the cells tuned to that direction to fire more strongly, the cells tuned to other directions will still fire in response to the incoherent motion noise (although cells tuned to the other direction will suppress their activity), and this summed activity may equal or outweigh the activity from the cells responding the coherent motion (Braddick et al., 2001). In addition, the motion sensitive complex in humans typically identified in MRI (MT+) is generally considered to be made up of a number of motion sensitive areas (hMT, hMST etc) whose cells may have different responses to coherent and incoherent motion (Becker et al., 2008).

The mixture of selectivites within a given voxel also makes studying the tuning of cells using fMRI difficult using standard univariate techniques. A common method of dealing with this is the use of adaptation techniques originally used in psychophysical studies, where the BOLD response to a probe stimulus is measured after prolonged exposure to an adapting stimulus, with a recovery from adaptation indicating that a separate population of neurons responds to the probe stimulus. Such methods have

been used previously to demonstrate tuning for orientation and direction of motion (Fang et al., 2005; Huk et al., 2001; Larsson and Heeger, 2006; Sapountzis et al., 2010). Although such methods are able to circumvent the resolution issues of fMRI, the exact choice of stimulus parameters can be a key issue (Fang et al., 2005; Larsson et al., 2006), which may require extensive piloting, and interpreting a positive adaptation result can be problematic when deciding which areas show selectivity (Bartels et al., 2008).

An alternative, and potentially more powerful method of demonstrating selectivity is the use of multivariate classification techniques. These methods have successfully been used to demonstrate selectivity for orientation and direction of motion (Kamitani and Tong, 2005, 2006), and in the case of orientation have been shown to be successful at discriminating differences in orientation finer than those possible with adaptation techniques (Sapountzis et al., 2010). In the next chapter I will examine multivariate classification of motion-direction in early visual cortex and MT+.

Chapter 4

Classification of Motion Direction

4.1 Classification with High Field fMRI

Multivariate pattern analysis has been shown to be able to 'decode' stimulus properties from patterns of BOLD activity in fMRI. For example, the perceived direction of motion for a cloud of drifting dots can be decoded from the patterns of activity in early visual cortex and in MT+ (Kamitani and Tong, 2006; Serences and Boynton, 2007b). This is despite the fact that the voxel sizes used, typically ~3 mm isotropic in standard fMRI experiments at 3T, are much larger than those required to individually sample the underlying neural architecture, as each voxel contains tens of thousands of neurons with a variety of different tuning properties. Voxels of this size are also too coarse to directly sample cortical columns, the smallest neural feature likely to be directly measurable using magnetic imaging methods, which are at scales of less than 1 mm in humans. This presents a question: how can a stimulus specific signal be present in the voxel sizes typically used in fMRI?

The reason that a single voxel normally lacks selectivity lies in the way it samples the underlying neural activity. A common way of thinking about a voxel is as a 'compact-kernel filter' (Kriegeskorte et al., 2010), where a voxel samples the neuronal activity pattern by averaging together the signal within its boundaries. To include the effects of hemodynamic blurring, a more accurate model has the voxel sampling the neuronal pattern via a Gaussian kernel. The voxel grid in this case can be thought of as creating a downsampled/blurred image of the neuronal activity pattern on the cortical surface. At standard resolutions, information at the scale of cortical columns should be too fine to resolve given the voxel sizes. The individual voxels would be unlikely to all contain identical proportions of neurons/columns selective for different orientations (or

other stimulus features), which could lead to slight biases for different stimuli for each voxel (Boynton, 2005). Although weak at the level of single voxels, these biases could be detected at the multivariate level by a classifier algorithm. This explanation represents a form of aliasing of fine-grained neuronal information into the coarser scale voxel pattern. The possibility that standard resolution fMRI could be made sensitive to fine-grained patterns of columnar activity at scales that cannot be represented directly in the voxel patterns (this being information in the 'hyper-band' frequencies of the neuronal activity pattern, which are those above the Nyquist limit of the voxel sampling (Kriegeskorte et al., 2010)) has been called 'fMRI-hyperacuity' (Op de Beeck, 2010).

Arguments against the hyperacuity hypothesis stem from the fact that blurring or downsampling of the voxel patterns prior to classification does not degrade performance of the classifiers. Op de Beeck (2010) blurred the voxel patterns in response to orientated gratings with a Gaussian kernel prior to classification, and found performance unaffected using kernels of up to 8 mm Full Width at Half Magnitude (FWHM). Gaussian blurring reduces the fine-detail in a pattern, acting as a low-pass spatial filter that emphasises information at coarser spatial scales, so successful classification after the reduction of fine-scale information was taken as indication that information at a coarser spatial scale must be used by the classifier. Smoothing simulated neuronal data sampled by a standard voxel grid, with information at either a fine (< 1 mm) or coarse (> 4 mm) scale showed unaffected performance in the latter case, but not the former, further indicating a role for a coarse signal in classification performance (Op de Beeck, 2010).

Evidence for a coarse scale signal was also found by Gardner et al. (2006), who found above-chance classification for orientation even when their voxels had been downsampled to an inplane resolution of 1 cm. Kriegeskorte et al. (2010) modelled classification results for stimulated eye based on ODCs and found that classification results should decline rapidly as voxel was size was increased, and that a larger voxel should be far less informative than the equivalent set of smaller voxels even when noise cancellation was taken into account. Results such as these were taken as evidence that the signal used by classifiers must lie at a much coarser scale than the fine-grained patterns of columnar activity.

Counter to these arguments, it has been suggested that Gaussian blurring and downsampling may not prove that only coarse-scale information is used. Kamitani and Sawahata (2010) argued that Gaussian blurring does not actually *remove* high-spatial frequency information in an image, but simply scales down both the signal and noise in the high-frequency bands of a pattern. If the kernel used to filter the image is known,

the process is invertible and the original image can be recovered, indicating the continued presence of detailed pattern information after smoothing. Therefore, they argued that the unaffected performance after Gaussian blurring did not indicate the use of solely coarse spatial scale information by the classifier.

Although results from smoothing/downsampling do not completely preclude fMRI hyperacuity via aliasing by voxels, it is unlikely to be the source of classifier performance for a number of reasons. 1) Any aliased signal would be fairly low in amplitude, due to the local averaging/Gaussian blurring taking place within each voxel acting as a kind of 'anti-aliasing' filter (Kriegeskorte et al., 2010), reducing the projection hyper-band into lower spatial-frequency bands. Although this low-pass filtering will not be complete, the aliased signal is expected to be very low in amplitude. 2) The effect of head-motion on such aliased signals (Kriegeskorte et al., 2010). Moving the voxel-grid by even a small amount will completely change the aliasing patterns of the hyper-band information, in a way that is not possible to correct via motion-correction algorithms. Given the successful classification results seen in fMRI studies, where some head-motion is unavoidable, and especially given the fact that classification can work between different scanning sessions (Kamitani and Tong, 2005, 2006), fMRI hyperacuity via aliasing seems unlikely. 3) Additionally, band limitations in phase- and frequency-encoding directions of MR imaging make aliasing of hyper-band information extremely unlikely (Swisher et al., 2010).

Whilst the aliasing of 'hyper-band' information into lower spatial frequency bands appears unlikely given a standard model of the voxel, columnar patterns are unlikely to have information solely at these spatial scales. Random variation in the spatial distribution of columns should lead to biases amongst different voxels, manifesting as components of the pattern at a range of different spatial-frequencies, including those directly measurable with fMRI (Kamitani and Tong, 2005). This information is not aliased hyper-band information, but rather reflects low band neuronal-pattern information. Direct study of orientation columns in the macaque has shown that weak, low spatial-frequency signals of this type exist (Obermayer and Blasdel, 1993).

High-resolution fMRI of cat visual cortex, at resolutions high enough to resolve individual orientation columns, followed by ideal high and low pass filtering (which offers greater spatial frequency isolation than Gaussian filtering) and multivariate classification analysis, showed that whilst classification was highest at the scale of the cortical columns themselves, the orientation signal was spatially broadband, extending to the scale of millimetres (Swisher et al., 2010). Low-pass filtering that obscured the orientation columns themselves still showed significant biases within voxels, with perfor-

mance remaining well above chance with filter sizes up to ~1 cm. High-pass filtering showed performance equivalent to the original images when pattern information at scales of ~1.3 mm were included, matching the average spacing between the orientation columns, with no increase in performance after the addition of coarser scale information.

A similar pattern of results was seen in high-resolution fMRI of human visual cortex, with low-pass and high-pass filtering followed by orientation classification showing a strong orientation signal at the scale of ~2-10 mm, consistent with a signal originating from variation in the spatial distribution of cortical columns (Swisher et al., 2010). Although this result was found using ideal volumetric filtering, which removes rather than attenuates unwanted spatial frequencies, when the same analysis was performed with a Gaussian filter similar results were found, although the drop-off in performance for low-pass filtering was less steep than for ideal-filtering. This contrasted with the findings of Op de Beeck (2010) who found no drop in performance up to 8 mm FWHM.

One possible explanation could be that Op de Beeck (2010) used 3 mm isotropic voxels, whereas Swisher et al. (2010) used 1 mm isotropic voxels. Swisher et al. (2010) reanalyzed data collected by Kamitani and Tong (2006) with 3 mm isotropic voxels, and found significant reduction in classification accuracy after 8 mm smoothing, again in contrast to the findings of Op de Beeck (2010). Swisher et al. (2010) hypothesized that the difference in findings may be due to differing levels of head-motion, as they found a single subject with a large amount of head-movement showed no loss in performance after 8 mm of blurring. Misaki et al. (2012) showed that the effects of smoothing or ideal filtering could vary from subject to subject, with some of the effects being due to head-movements, but some variability existing beyond this that could not be accounted for by head-motion (Misaki et al., 2012).

Information at even coarser spatial scales has also been suggested as a potential source of classification performance, based on the findings of above chance performance after blurring (Op de Beeck, 2010) or subsampling (Gardner et al., 2006) prior to classification. Some potential coarse-scale biases that could be used by a classifier are global biases for certain stimuli types, for example biases for cardinal (Furmanski and Engel, 2000) or radial orientations, the latter of which has been demonstrated electrophysiologically in animals (Leventhal, 1983), and in humans and primates using fMRI (Freeman et al., 2011; Sasaki et al., 2006). A weak but reliable bias amongst voxels for radial orientations could be used by multivariate methods to decode orientation, without the requirement of signals arising from local variations in columnar distribution. Initial studies of orientation classification ruled out a role for a radial bias for orientation

based on the finding that removing the radial component of the pattern before classification did not harm performance (Harrison and Tong, 2009; Kamitani and Tong, 2005). Swisher et al. (2010) found some evidence of a radial bias for orientation, albeit an extremely weak one seen only when the data had been low-pass filtered extensively, and they concluded that any radial bias provided little information for a classifier.

Evidence that orientation decoding may rely on larger scale biases came from Freeman et al. (2011), who argued that an analysis based on spatial filtering assumes that signals arising from a radial bias would appear exclusively at low spatial frequencies and columnar signals at higher frequencies. They demonstrated that this reasoning may not hold by showing that decoding of information known to have coarse scale organization and expected to have no contributions from columnar irregularities, angular position in a retinotopic map, also showed a detriment in classification after low-pass filtering. Using the same high- and low-pass filtering methods as Swisher et al. (2010), they showed the same pattern of results for orientation and angular-position classification, with an drop in performance after low-pass filtering and classification well above chance after high-pass filtering. Therefore they concluded that responses organized into a coarse map would lead to signals at multiple spatial scales, so impaired performance after low-pass filtering could not rule out a coarse-scale map driving classification performance. They additionally found that averaging voxels together in a way that preserved a bias for radial orientations, by averaging voxels together based on the independently measured representation of polar angle, showed preserved classification performance (Freeman et al., 2011). When the equivalent level of averaging was performed randomly, classification performance was strongly reduced. Removing the radial bias via projection reduced classification accuracy, leading them to conclude that a radial bias organized at a very coarse scale was both *necessary* and *sufficient* for successful orientation classification.

The necessity for a radial bias for classification has been questioned, given that orientation stimuli without radial biases can be successfully classified using MVPA (Clifford et al., published on-line 26 April 2011; Hong et al., 2011; Kamitani and Tong, 2006; Mannion et al., 2009). Clifford et al. (published on-line 26 April 2011) hypothesised that the stimuli used in Freeman et al. (2011), a rotating grating stimulus similar to the rotating wedge stimulus used in retinotopy, may have been particularly dependent on a coarse scale radial bias. Classifying 'block design' stimuli', rather than the 'event related' style stimulus used by Freeman et al. (2011), may be more dependent on columnar information at a range of spatial frequencies.

In addition to neural pattern information at various scales, it has been suggested that

the neural vasculature itself may have some selectivity, based on it how it samples the underlying cortical maps (Kamitani and Tong, 2005). If draining veins sample selective columns non-homogeneously for example, they will attain some selectivity that can be imaged directly by the voxels, and it may be the case that vasculature is in fact organized around a columnar arrangement in a way that amplifies weak selectivity signals (Gardner, 2010). Empirical evidence for a selective signal arising from draining veins came from a high-resolution classification study that also collected T_2* weighted venograms (Gardner et al., 2006), demonstrating large draining veins with orientation specific responses that could support successful classification even in heavily down-sampled images. A study of decoding which eye was stimulated from patterns of responses in V1 at high-resolution also showed a contribution to decoding from draining veins that sample ODCs in a biased fashion (Shmuel et al., 2010). Selective signals were also found in voxels completely corresponding to grey matter at both high and low resolutions, possibly due to local variations in pattern of ODCs, indicating that venous signals were not the exclusive source of classification performance (Shmuel et al., 2007). However these studies were performed at a high magnetic field strength (7T) compared to earlier classification studies. Imaging at 7T is known to reduce the intravascular signal from veins compared to tissue, so it may be the case that venous signals are more dominant at standard field strengths (Shmuel et al., 2010).

It may also be the case that the vasculature allows selectivity even without an explicit biased sampling of neuronal selectivity. Under this model the fine structure of the vasculature supplying the BOLD signal to each voxel causes that voxel to act as a 'complex spatio-temporal' filter of the underlying neural activity pattern, rather than as a 'compact-kernel' filter (Kriegeskorte et al., 2010). The aliasing of neuronal patterns at different spatial scales into the different spatial frequency bands is expected to be more complex than expected by a simple averaging model, and as the aliasing is done by the vasculature rather than the voxels themselves, the aliased signal will be less sensitive to shifts of the voxel grid due to head movement. Supporting this hypothesis, Kamitani and Tong (2005, 2006) were able to show classification between scanning sessions for individual subjects, where the sampling of neurons/columns by the voxel grid is very unlikely to the same from session to session, due to the inability to exactly replicate the position of the voxel sampling grid with regards to the subject.

The exact contribution from the neuronal pattern as sampled by the voxels, whether via local variations captured directly by voxels or a coarse scale map, and the signal sampled via the vasculature itself, whether due to biased sampling or a complex spatio-temporal filter, to classification performance is an ongoing area of debate (Kriegeskorte

et al., 2010). A study modelling classification accuracies of stimulated eye via ODCs based on a 'compact-kernel' model of voxel sampling concluded that whilst hyperacuity could be ruled out, local variations in the pattern of ODCs that could be captured by the voxels were sufficient for above chance classification (Chaimow et al., 2011). However, the classification accuracies they found in the model were less than those typically found in classification experiments, leading them to conclude that coarse-scale signals and some form of vascular contribution, be it biased sampling or complex spatio-temporal filter, must contribute the additional selective signal. Whilst neuronal pattern information, whether in the hyper-band or at spatial scales directly measurable by the voxels, and vascular selectivity are all ultimately related to neuronal activity, and therefore measures of neuronal selectivity, the exact localisation of the signal used by the classifier can be important if strong conclusions about the underlying neural signal are to be drawn. Therefore establishing the scale of the information used by classifiers is important if we wish to draw conclusions about the neural architecture in a given area.

Previous analyses of motion selectivity have found that the perceived direction of motion of a drifting cloud of dots could be decoded from patterns of fMRI activity in both early visual cortex and MT+ in humans (Kamitani and Tong, 2006; Serences and Boynton, 2007b). Classification was well above chance for data from all areas, with no greater classification accuracy in MT+ compared to V1, despite this area's assumed high proportion of direction selective neurons. Based on these previous reports, we attempted to replicate this result using higher-resolution fMRI at 7T, to see if we could draw any conclusions about the scale of the signal used in the classification of direction of motion. Moving to this higher field allowed the voxel size to be reduced from the 3 mm isotropic voxels used in the original study to 1.5 mm isotropic voxels, an eight-fold reduction in volume. If successful classification is in fact based on the distribution of columns between voxels, then decreasing the size of the voxels should increase the amount of information available to the classifier by increasing the biases available present in each voxel (Kriegeskorte et al., 2010). However, if information is in fact carried by a signal at a coarser scale, then increasing the resolution may be likely to have little to no effect on performance (Gardner et al., 2006; Op de Beeck, 2010). Additionally, although the voxels in the current experiment are slightly larger than those used by Swisher et al. (2010) (1 mm^3 versus 3.375 mm^3), using smaller voxels than those used in standard fMRI experiments allows us to investigate contributions of variations in the spatial pattern at a scale smaller than these voxel sizes to classification.

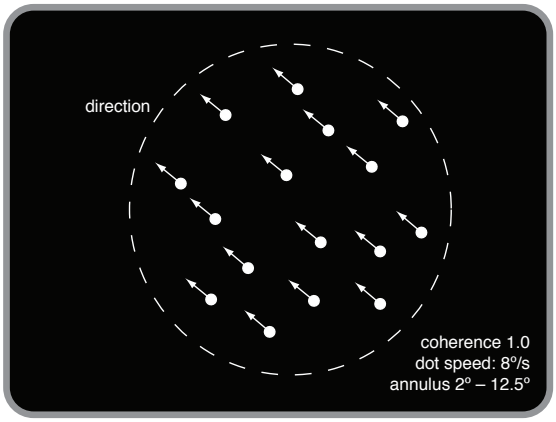

Figure 4.1: Examples of stimuli used in motion classification scans. All the dots in the field drifted in a common direction (white arrow), which changed every 16s. Note that the dot density and size are not to scale in this illustration. In the actual stimulus display, dot density was 5 dots per /deg^2 and dot size was 0.1 deg

4.1.1 Methods

Subjects

Five subjects experienced in fMRI experiments and with normal vision participated in this study with written consent. Procedures were conducted with approval from the University of Nottingham Medical School ethics committee.

Stimuli and Task

Two experimental paradigms were used in each scan session: 1) A minimum of two scans of the motion retinotopy experiment detailed in Section 2.2, and 2) a block paradigm with motion in one of eight possible directions. The block-classification stimuli consisted of white dots in a circular aperture (diameter 24.6° of visual angle) (Figure 4.1). For these stimuli, all the dots drifted in the same direction at the same time for blocks of 16s (order randomized), following the paradigm of (Kamitani and Tong, 2006), with a block of fixation at the beginning and end of each scan (total, 2 minutes 40 seconds per scan). The dots had a lifetime of 200 ms. Subjects completed 16-20 of the classification runs per scan-session. To control for the effects of changes in attentional state, subjects performed a contrast discrimination task at fixation (as detailed in Section 2.3.4).

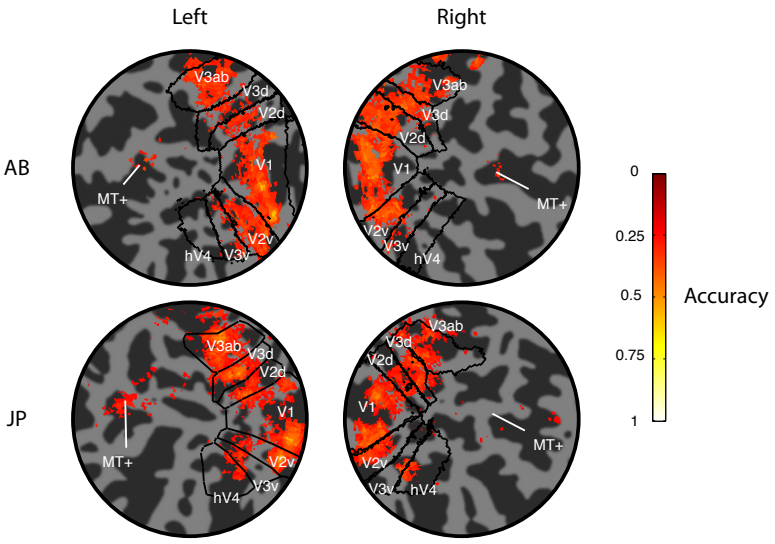

Figure 4.2: Classification accuracy maps derived from a searchlight classification analysis using a spherical searchlight with a radius of 3 voxels. Voxels are thresholded on the basis of a binomial test ($p<.05$), corrected for multiple comparisons (FDR).

4.1.2 Results

Figure 4.2 shows the results of a searchlight analysis for two subjects, using a spherical searchlight with a radius of 3 voxels. Accuracies are thresholded for significance using a binomial test, corrected for multiple comparisons (False Discovery Rate). Elevated accuracies are seen across early visual cortex, particularly in more dorsal areas. Some significant accuracies are also seen in an area corresponding to MT+. This replicates previous findings of selectivity for direction of motion seen in these areas in previous studies.

In the ROI-based multivariate analysis, we found above-chance classification accuracies for direction of motion classification in areas V1-V4 and MT+/V5 for all 5 subjects (Figure 4.3). Permutation tests confirmed this for each ROI for each individual subject. As our experiment was based on 8 possible directions of motion, the chance level for *proportion correct* was 0.125. In line with previous results, but perhaps surprisingly given previous single cell studies, classification accuracy was higher in earlier visual areas (V1-V3) than in area MT+/V5 (Figure 4.3), which is known to show strong selectivity for motion at the single-cell level (Snowden et al., 1992).

Previous studies hypothesised the reduced number of voxels available for classification as a reason for the reduced accuracy seen in MT+, as similar levels of classification

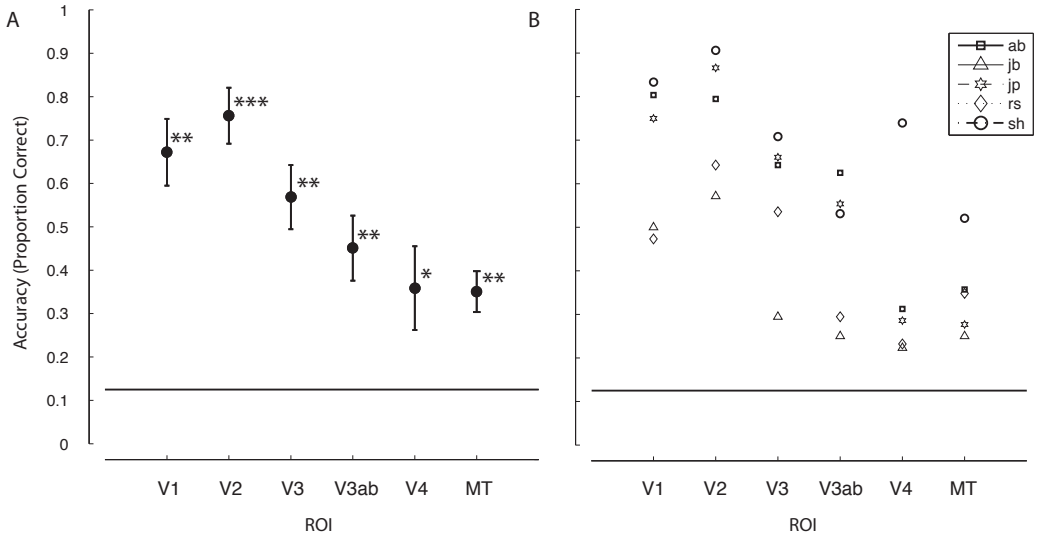

Figure 4.3: A) Mean classification accuracies of 8-way direction classification in retinotopically defined ROIs. Error bars, standard error across 5 subjects. Horizontal line indicates chance performance, p(correct)=0.125. Asterisks indicate significance on a one-sample t-test (* p < .05; ** p < .01; *** p < .001; one-tailed) B) Data plotted for individual subjects.

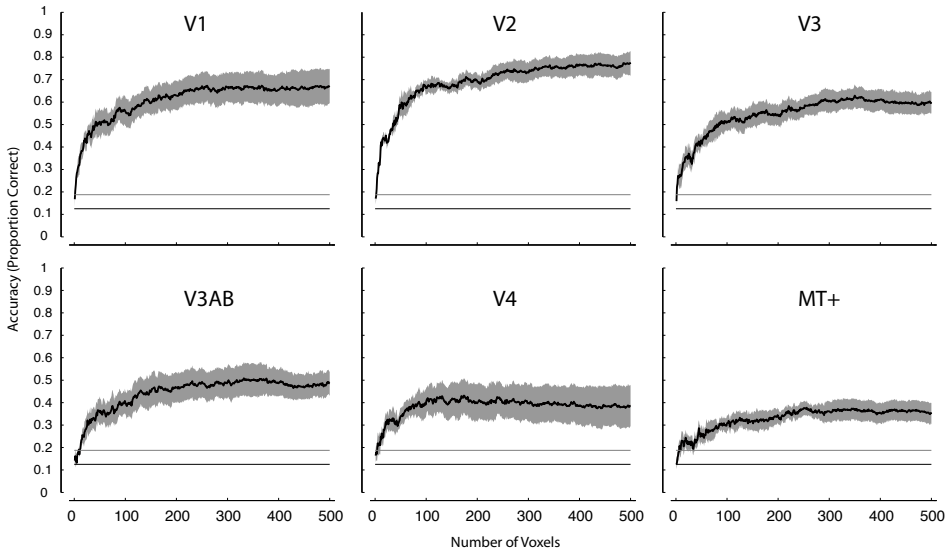

Figure 4.4: Classification as a function of number of voxels. Voxels were sorted according to their response to the retinotopic localiser. Black line indicates chance performance, p(correct)=0.125. Grey line indicates 95th percentile of null distribution. Error surface = S.E.M across 5 subjects.

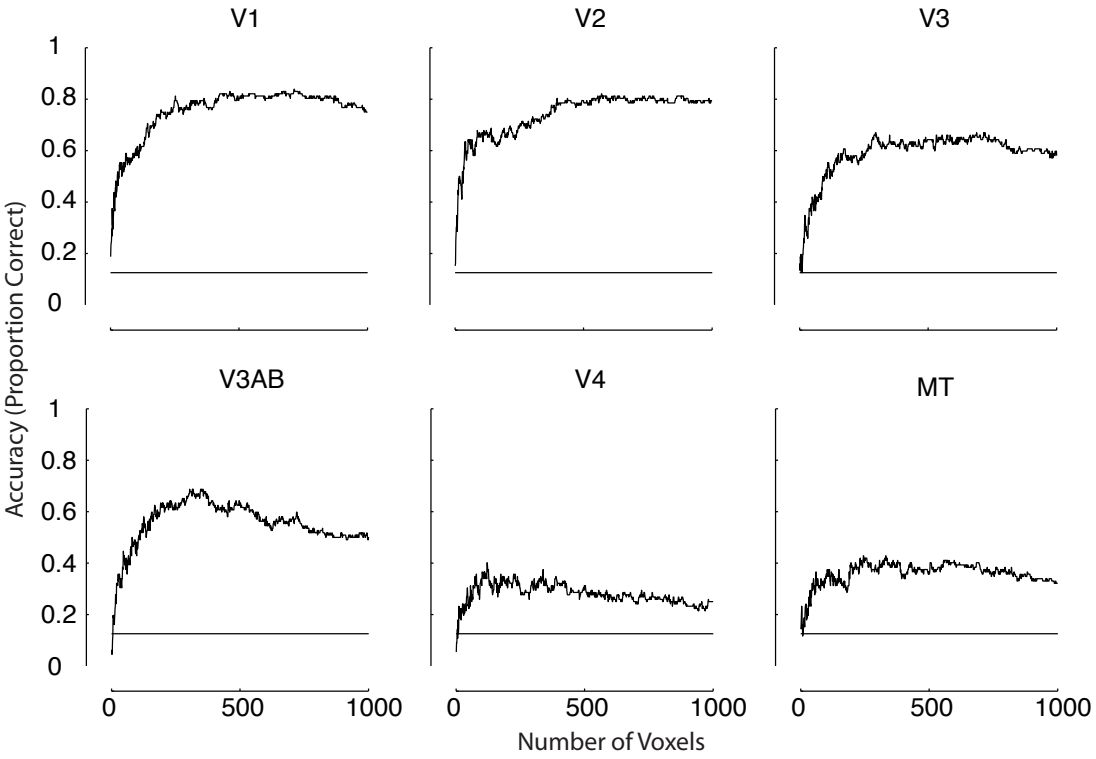

Figure 4.5: Classification as a function of number of voxels for a single subject. Early visual areas show asymptotic performance beyond 500 voxels, whereas later areas show a dip in performance beyond this point. Voxels were sorted according to their response to the retinotopic localiser.

accuracy were seen when the number of voxels in V1-V4 was reduced from 200 to 100 to match MT+ (Kamitani and Tong, 2006). Using the higher resolution afforded by 7T fMRI, the number of voxels from each ROI in the current experiment was matched at 500, creating equal voxel numbers in all studied ROIs. Classification accuracy was still lower in MT+ compared to earlier visual areas even with this increased number of voxels. Reducing the number of voxels below 500 reduced classification accuracy in all areas (Figure 4.4), with a plateau at around 200-300 voxels. Looking at the results from a subject who had 1000 voxels available from all ROIs (Figure 4.5), no improvements were seen when the number voxels was increased from 500 to 1000, with some areas showing a decline in performance. Therefore 500 voxels was chosen as a value that maximised the amount of features available for classification, whilst allowing comparison between visual areas.

The classification results can also be used to construct 'tuning-curves' for each direction, by showing how often each direction was classified as one of the 8 possible direc-

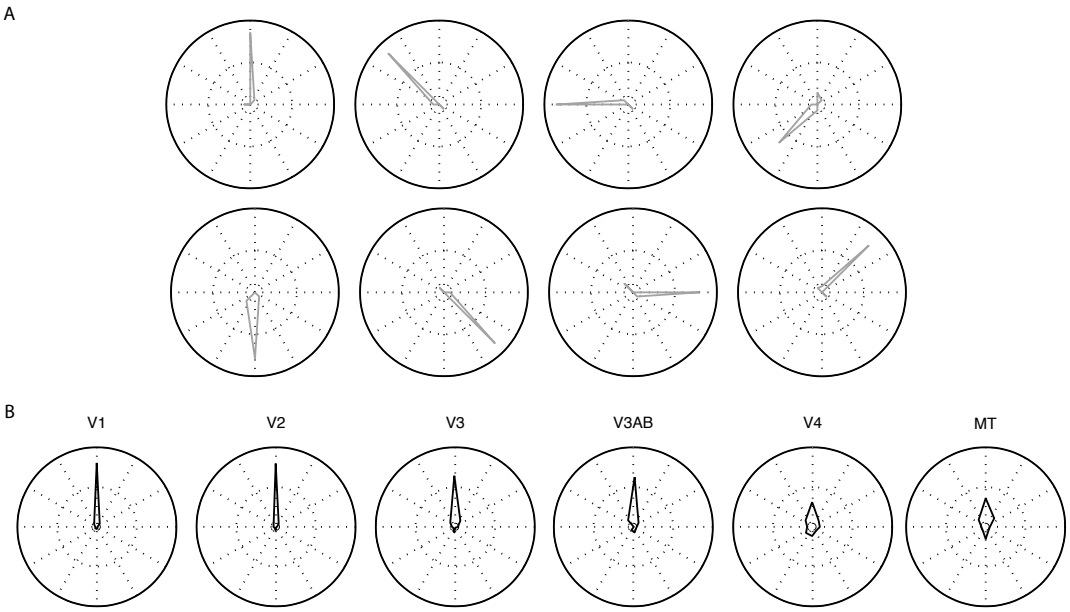

Figure 4.6: Tuning curves (distribution of decoded directions) for multivariate classification for a single subject. Polar angles indicates classified stimulus motion direction in degrees, and distance from origin indicates proportion of stimuli classified as that direction. Top Row: Tuning Curves for each of the 8 possible direction stimuli from V1. Bottom Row: Combined tuning curves (collapsed across direction) for each ROI.

tions (Kamitani and Tong, 2006). Tuning curves for one subject are shown in Figure 4.6, Top Row. In V1, the majority of directions were classified correctly, leading to a peak in each tuning curve. The most common misclassifications were for directions either side of the actual direction, and very few misclassifications were for orthogonal or opposite directions, creating a very narrow tuning curve for each direction. By recentering each tuning curve to a direction 0°, the tuning curves for multiple directions can be combined, yielding a single tuning curve for each ROI (Figure 4.6, Bottom Row). The tuning curves for early visual areas are fairly similar, growing slightly broader in the higher areas. The curve for MT+ is considerably broader than the earlier visual areas, matching earlier results at 3T (Kamitani and Tong, 2006).

An alternative method of analysing data from multiple voxels are *voxel tuning curves* (Serences et al., 2009), which allow the tuning functions of the voxels themselves to be studied, as opposed to the tuning of the classifiers. Voxels were assigned to bins based on the direction of motion that elicited the largest mean response over the same period used to create the patterns for classification, after removing the mean response across all voxels at each direction to correct for main effects that affected each voxel

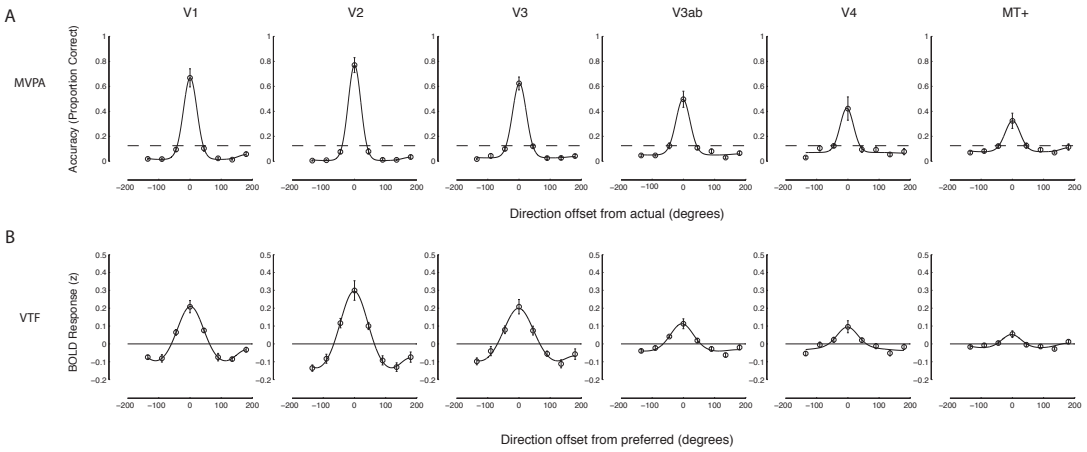

Figure 4.7: Mean Tuning curves for all subjects. A: Symbols, MVPA Tuning Curves; line, circular Gaussian fit. B: Symbols, Voxel Tuning Functions; line, circular Gaussian fit. Error bars = ± 1 SEM across subjects (n=5).

response. Voxels were binned based on the data from all but one scan, and then the response of the voxels in each bin were computed using data from the remaining scan. Data were then combined to form an average tuning function for each ROI. The VTF data showed a similar pattern as the MVPA-tuning functions, with directions either side of the preferred direction being the next most active. This indicates that patterns of BOLD activity for adjacent directions are similar, and that voxels show a preference for a given direction. An elevation in accuracy/response was also seen at 180° relative to the actual/preferred direction in some visual areas.

Figure 4.7 shows the data from the MVPA and VTF analyses of the data, along with a least squares fit using the sum of two Circular Gaussian (Von Mises) functions of the form:

$$\delta = b + a * e^{[k*cos(x-\mu)-1]} \tag{4.1.1}$$

where a is the response amplitude, b is the baseline, and k is the concentration parameter (Serences et al., 2009). The second Gaussian function's peak was constrained to be 180° separated from the peak of the first Gaussian function in order to capture the second peak relative to the preferred/correct direction.

To assess whether increasing the resolution increased the accuracy of the classifier, we sub-sampled our data by a factor of 2. This increased the voxel dimensions from 1.5 mm isotropic to 3 mm isotropic, matching the dimensions of the original Kamitani and Tong (2006) experiment. Overall, we found little to no reduction in performance in V1

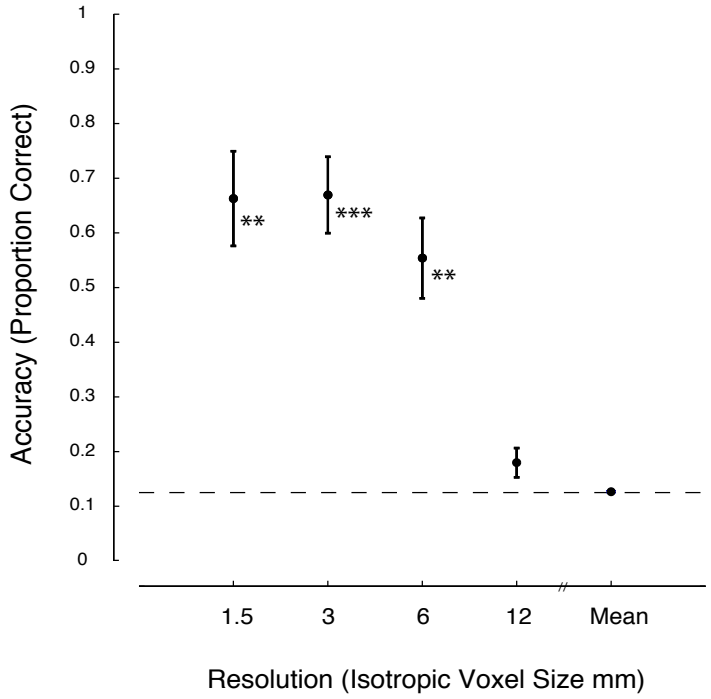

Figure 4.8: Effect of re-sampling the data on classification accuracy in V1. Re-sampling the 1.5 mm isotropic voxels to 'standard' resolution (3 mm isotropic) showed no drop in performance. Re-sampling further to 6 mm isotropic (a 64 fold increase in voxel volume relative to the original voxels) showed a significant drop in performance, although performance was still significantly above chance. Re-sampling further to 12 mm isotropic caused performance to drop to chance, equivalent to using the mean time-series of the ROI. Black Asterisks show significance in a one-tailed t-test testing for greater than chance accuracy.

when this reduction was made. Further sub-sampling, increasing the voxel by a factor of 2 to 6 mm isotropic showed a drop in performance, although still significantly above chance (one tailed $t[4]=5.8351$, $P = 0.0021$) and dropping to 12 mm isotropic showed near chance performance (Figure 4.8), equivalent to using the mean time-series from the ROI.

An additional method of assessing the scale of the signal used by the classifier is blurring the data prior to classification (Op de Beeck, 2010; Swisher et al., 2010). This manipulation would be expected to remove any fine-detail information and leave only coarse-scale information. If classification performance persists after blurring, this could indicate that fine-scale information is not used by the classifier. We tested the effect of such blurring on motion classification in V1 via low- and high-pass filtering with Gaus-

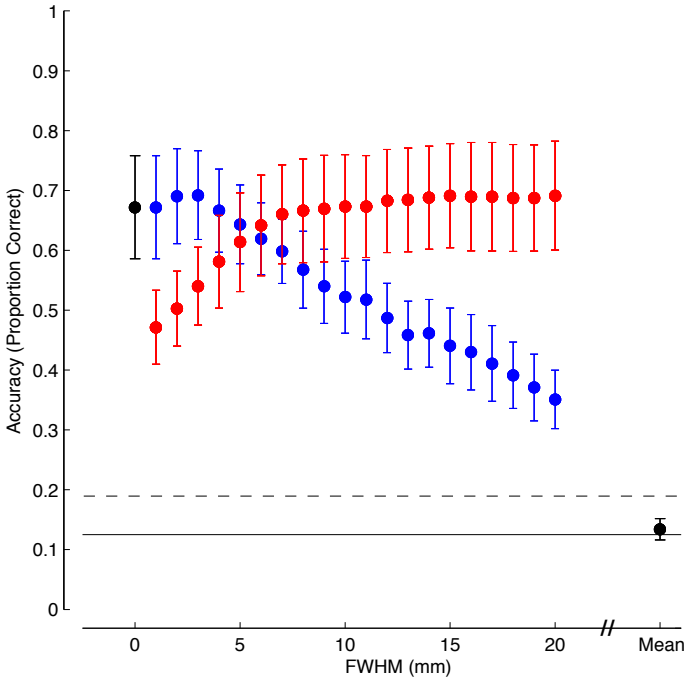

Figure 4.9: Effect of filtering the data on classification accuracy in V1. Blue points show the effect of low-pass filtering, red points show the effect of equivalent high-pass filtering. Black point shows the classification accuracy using the mean time-series of the ROI. Black line indicates chance performance, p(correct)=0.125, dotted line = 95th percentile of null distribution.

sian kernels of varying Full-Width at Half-Maximum (FWHM). Low-pass filtering was achieved using the fslmaths tool available as part of the FMRIB Software Library (FSL), and equivalent high-pass filtering was done by subtracting the low-pass image from the original unfiltered image. Low-pass filtering led to a slight increase in accuracy, followed by a steady decline but remaining above-chance even up to a FWHM of 2 cm (Figure 4.9, Blue Symbols). Using high-pass filtering, classification returned to initial levels when a Gaussian kernel with a FWHM greater than ~5 mm was used (Figure 4.9, Red Symbols). Classification using the mean time-series of the whole ROI was at chance (Figure 4.9, Black Point).

4.1.3 Discussion

Using stimuli and methods originally used by Kamitani and Tong (2006), we found above chance classification of the observed motion of drifting dot patterns in areas V1-V4 and MT+/V5, with higher levels of accuracy in the earlier visual areas. The finding that area MT+/V5 shows lower classification accuracy than earlier visual ar-

eas, despite macaque MT+/V5 containing a greater proportion of direction selective cells, was suggested by Kamitani and Tong (2006) to arise from the smaller number of voxels available from MT+/V5 compared to V1, as when they reduced the number of voxels used in early visual areas, classification accuracies were equivalent. However, given the higher resolution available at 7T, we were able to increase the number of voxels available for classification from MT+/V5, without improving the classification accuracy. This suggests that a simple reduced number of features is not the cause of this difference in performance. Other factors that could lead to reduced performance even if equivalent (or greater) selectivity exists in MT+/V5 are the arrangement or distribution of direction selective columns leading to very small voxel-wise biases (Kamitani and Tong, 2006; Malonek et al., 1994), or differences in the amplitude of BOLD response in different parts of the cortex (Smith et al., 2010; Tong et al., 2012). Due to the wide range of possible factors involved in determining classification accuracy in a given area, it is difficult to compare classification performance across cortical areas and inferring differences in selectivity from classification results.

The distribution of classified directions for each direction (MVPA tuning functions) for a single subject is shown in the top row of Figure 4.6. There is a peak at the correct direction (indicating that majority of stimuli were classified correctly). If directions are misclassified, they tend to be classified as directions either side of correct direction, with no directions misclassified as the orthogonal directions of motion. The bottom row of Figure 4.6 shows the average tuning curves for V1-V4 and MT+ for a single subject. The same pattern is seen across all visual areas, with a peak at the correct direction, the majority of misclassifications occurring either side of the correct direction, and few or no misclassifications for orthogonal directions.

The average MVPA tuning functions for each ROI across all the subjects are shown in Figure 4.7 A, along with a a fitted curve consisting of the sum of two Circular Gaussian functions. Results matched those in Kamitani and Tong (2006), with a peak at the actual direction, the next most common classifications being either side of the actual, and very few misclassifications at other directions. This indicates that the similarity of the activation patterns is determined by the similarity of the stimuli, so stimuli close to each other in terms of direction are more likely to be confused by the classifier. A small peak was seen opposite to the actual direction, indicating that stimuli were more often misidentified as the opposite direction, and rarely as an orthogonal direction. This mirrors a finding by Kamitani and Tong (2006), who also found that stimuli were more often misclassified in the opposite rather than orthogonal direction. They hypothesized this may be due to the columnar organization on the cortex, which in some animals

is known to abruptly shift preference by 180° (Shmuel and Grinvald, 1996). If this pattern persists in human visual cortex, this could mean that columns with opposite directional preference are more likely to be sampled in the same voxel, leading to the misclassifications seen here.

We also constructed VTFs for each ROI and subject using the methods set out by Serences et al. (2009) (Figure 4.7 B), again fitting two Circular Gaussians to the data to model the tuning curve. The pattern in the VTF results mirrors that of the MVPA results. The preferred direction for each voxel bin shows the highest amplitude response, the adjacent directions, the next highest, the lowest responses for the orthogonal directions, and second peak at the anti-preferred direction. As with the MVPA tuning, the second peak diminished in amplitude for area MT. Kamitani and Tong (2006) reported that none of their voxels showed a strong preference for a particular direction of motion, which appears to run counter to our results. Serences et al. (2009) repeated the Kamitani and Tong (2005) orientation study at a higher resolution (2 x 2 x 1.5 mm compared to 3 mm isotropic), and found proportionally higher amplitude VTFs than those seen in the original study, where voxels only showed very slight biases for their preferred direction. They hypothesised this may be due to their increase in resolution isolating the orientation columns more effectively, with less partial voluming effects caused by white matter being included within a voxel.

One key question for the current experiment was the effect of resolution on classification performance: does the use of high-resolution imaging allow greater classification accuracy than standard fMRI resolutions. If it is in fact the distribution of selective columns between voxels that leads to the slight biases between voxels that allows MVPA classification, then using smaller voxels might be expected to increase these biases, and allow more accurate classification (Kriegeskorte et al., 2010). Re-sampling the data prior to classification to match the resolution of earlier studies did not reduce the classification accuracies found with 1.5 mm isotropic voxels, and above chance classification was even found when voxels were further re-sampled to 6 mm isotropic resolution, despite this being 64 fold increase in voxel size. This is in line with previous studies showing that Gaussian blurring of the voxel patterns prior to classification analysis (Op de Beeck, 2010) or down-sampling data to lower resolution after acquisition (Gardner et al., 2006) does not harm performance, suggesting that fine detail information is not the source of the classification. There is some debate as to how much of the fine detail information is removed by Gaussian blurring. Kamitani and Sawahata (2010) argued that blurring/smoothing, specifically convolution of the data with a Gaussian kernel, was an invertible transformation, i.e. that the original pattern could

be recovered from the smoothed one. This indicates that the fine-detail information was not actually lost in the case of smoothing, and therefore successful classification after Gaussian smoothing does not tell us anything about the scale of the information used. However, Misaki et al. (2012) argued that the use of leave-one-run-out cross validation would in fact mean that classification could be made sensitive to even invertible transformations (such as Gaussian smoothing), such that classification performance after varying levels of Gaussian smoothing was informative about the scale of the signal used. Down-sampling the patterns, rather than blurring them, does not preserve fine detail information, although the effect of noise-cancellation with larger voxels may counteract any loss of fine detail information (Kamitani and Sawahata, 2010). However, when comparing a large voxel to the equivalent smaller voxels that would sample the same space, the smaller voxels should be expected to carry more information Kriegeskorte et al. (2010). If instead the signal used by the classifier exists at a range of spatial scales, including those directly measurable by standard fMRI resolution, then an increase in the resolution of the voxels may not be expected to improve performance.

A previous study of orientation classification at high-resolution also investigated the scale of signal used by classifiers (Swisher et al., 2010). They collected data at 1 mm isotropic resolution, and then tested classification after both low and high-pass filtering. Contrary to the preserved performance after blurring found by Op de Beeck (2010), they found that low-pass filtering, which should remove fine-scale information, reduced classification performance. For equivalent high-pass filtering, performance was at chance if only very-fine spatial information was included. If patterns at the scale of ~2 mm or lower were included, classification was above chance, and performance returned to original levels as lower-spatial frequency patterns up to ~10 mm were reintroduced. From these results, (Swisher et al., 2010) concluded that the signals primarily used by multivariate classifiers were on the scale of millimetres, based on relatively low-spatial frequency variations in the columnar map, as suggested by Kamitani and Tong (2005). We performed the same analysis of our motion-classification in V1 by Gaussian-filtering the EPI images prior to classification. We found similar results to those found by Swisher et al. (2010), in that low-pass filtering with increasing Gaussian FWHM filtering led to a decline in performance, whereas high-pass filtering with larger FWHM larger than ~5 mm led to equivalent performance to unfiltered data (Figure 4.9). Although performance remained well above chance for low-pass Gaussian filtering up to 2 cm FWHM, this could be explained by our use of Gaussian filtering as opposed to ideal filtering, which attenuates rather than removes unwanted spatial frequencies. The same is true for the high-pass filtering, which is why high-pass filtering with a very small Gaussian kernel still yields above-chance performance. Although the

trend for increasing levels of low-pass filtering was a steady reduction in performance, there was a slight increase in performance after a small amount of low-pass filtering (~3 mm FWHM). Similar results were seen in the data of Swisher et al. (2010), which they interpreted as being due to the influences of head-motion on the disruption of fine-scale pattern information.

Swisher et al. (2010) interpreted their results as indicating that the signal utilised in MVPA was at the scale of millimetres, with little to no contribution from a coarse, centimetre scale signal (such as a global bias for cardinal or radial orientations or directions). However, Freeman et al. (2011) showed that high- and low-pass spatial filtering had equivalent effects of classification of orientation and retinotopic position, the latter of which is known to have a coarse scale organisation in visual cortex, indicating that a coarse-scale cortical signal may not be straightforwardly represented as a coarse-scale component in an EPI image. Classifier performance after high- or low-pass spatial filtering may not be a reliable indicator of the scale at which the information utilised by the classifier originates.

The reduced classification accuracy in MT+ when compared to earlier visual areas is perhaps surprising, given the fact that this area is thought to be highly direction selective. Our increased resolution allowed us to increase the number of voxels available to the classifier without an increase in classification accuracy, suggesting that a simple lack of features is not the reason for the apparent lack of selectivity in MT+. Another possible explanation is that the spatial arrangement of direction selective columns in MT+ may be much more uniform than in early visual areas (Malonek et al., 1994), meaning that there are no local variations for the voxel grid to sample such that each voxel shows only a very weak bias for a particular direction. Thus the reduced classification performance in MT+ compared to earlier areas may not reflect the responses of the underlying neurons, but simply their distribution relative to the voxel sampling grid (Bartels et al., 2008; Kamitani and Tong, 2006).

Another potential explanation for the higher classification accuracies seen in early visual cortex relative to MT+ for motion stimuli is the contribution of oriented 'motion-streaks', visual artefacts parallel to the trajectory of a fast-moving object caused by temporal integration (Apthorp et al., 2013; Bartels et al., 2008). Apthorp et al. (2013) showed that classification could generalize between patterns elicited by fast-moving dot stimuli and patterns elicited by orientation stimuli with a parallel orientation in early visual cortex, with no generalization between slow-moving dots and static orientation stimuli. They concluded that discrimination of non-opposing motion directions in early visual areas was based on the orientation of the elicited motion streaks. The

presence of motion streaks in fast-moving random dot stimuli was therefore suggested as a contributing factor for the relatively low classification accuracies seen in MT+ relative to early visual areas (Apthorp et al., 2013). Given that data from MT+ does not support successful orientation classification (Kamitani and Tong, 2005), oriented motion streaks are not expected to offer any benefit in the classification of non-opposing motion. Signals from early visual areas do allow concomitant orientation classification, and it has been suggested that the higher accuracies seen for motion classification in these areas is primarily driven by orientation information. The stimuli used in the current experiment were above the speed threshold for the creation of motion streaks (~2 deg/sec, Geisler (1999)), so the contribution of orientation information to the classification of motion-direction cannot be assessed.

Another potential contribution to classification performance is the possible presence or lack of a global bias for particular directions, for example a bias for cardinal or radial directions in different visual areas (Clifford et al., 2009; Raemaekers et al., 2009; Sasaki, 2007; Sasaki et al., 2006). Freeman et al. (2011) showed that a weak radial bias for orientation exists in visual cortex, and demonstrated that this bias was both sufficient and necessary for above chance classification. However, the necessity of a radial bias has been questioned, given that classification is possible for stimuli with balanced radial components (Hong et al., 2011; Kamitani and Tong, 2006; Mannion et al., 2009), and that classification is still possible when radial biases are accounted for (Kamitani and Tong, 2005, 2006).

4.2 Investigating the contribution of radial bias

There is evidence that multivariate classification could also be driven by coarse-scale information. One such potential source is a global bias for cardinal or radial orientations or directions (Clifford et al., 2009; Raemaekers et al., 2009; Sasaki, 2007; Sasaki et al., 2006). For example, in the orientation domain it has been demonstrated that a bias for radial orientations co-varies with retinotopy, that classification remains above-chance even after averaging voxels together based on their retinotopy, and that removing the map-based component of the orientation signal by projecting it out of the data reduces classification accuracy (Freeman et al., 2011). As discussed in Section 4.1, these results were suggested to demonstrate both the sufficiency and necessity of a retinotopically-organised preference for radial orientation for successful classification of orientation. Chaimow et al. (2011) suggested that information at this scale was the only potential source of the very high classification results seen in previous studies (if a compact-kernel sampling method is assumed). A number of previous classification studies however showed no reduction in performance after removing radial biases for orientation and motion on classification accuracy (Kamitani and Tong, 2005, 2006), and there is therefore still a debate about this question in the literature.

Several studies have demonstrated radial biases for motion in the visual system. Using random dot motion stimuli where the direction of global motion rotated by 360 degrees over a given period, Raemaekers et al. (2009) demonstrated a radial bias for motion in areas V1-V3, but not in MT+. Clifford et al. (2009) also showed a radial bias for motion in areas V1-V3 using motion defined contours of varying orientation, with the bias for radial motion being independent of the orientation of the contour, although that study did not consider area MT+. To test whether the coarse scale signal suggested by these experiments may be responsible for the classification seen in previous studies, in the current study we tested the effect of averaging based on voxels' retinotopic preference to examine the sufficiency of such a coarse scale preference for radial direction for successful classification.

We performed a further analysis of our motion classification data by undertaking a binning and averaging regime for the voxels in the various ROIs in this experiment.

4.2.1 Methods

To test whether a coarse-scale retinotopic preference for direction of motion, for example a preference for radial motion, was sufficient for direction classification, the classification analysis was repeated after averaging the input voxels based on the phase of

their response to the retinotopic stimulus. Voxels were assigned to bins, each bin corresponding to a separate range of polar angles. The time-series of the voxels within each bin were then averaged to yield a smaller number of 'super-voxels' (Freeman et al., 2011), which were then used to generate the patterns for classification analysis detailed above. This process was repeated for bins of varying width. To test whether this retinotopic scale signal allowed successful classification, we repeated the averaging process, but shuffled the phase values for each voxel prior to binning and averaging, leading to equivalent levels of averaging that did not preserve the retinotopic signal.

4.2.2 Results

If there is a bias in the voxels' preferences for direction of motions, such that preferred directions of motion are correlated with particular visual field angle preferences, then the details of the averaging regime should have an effect. In particular, we would expect classification following averaging that respects the retinotopic maps to be less disrupted. We found that for both averaging methods, decreasing the number of bins used (i.e. increasing the range of polar angles combined) led to a decrease in classification performance (see Figure 4.10). However, for the averaging scheme in which we respected the preferred visual field location for voxels – in effect preserving any information at the level of a coarse-scale map – performance dropped off less quickly. Classification performance remained above chance until the width of the bins was considerably larger than in the random averaging scheme (Figure 4.10). For example, in V1 the retinotopic binning method was at 0.5 accuracy at a bin width of ~30°, while the random method had equivalent accuracy with an equivalent width of ~3○. Therefore there was a roughly tenfold increase in the number of voxels (as judged by the amount of polar angle represented) that could be averaged together under the retinotopic binning method to yield an equivalent level of performance. This preserved performance in the retinotopic condition is represented by a rightward shift along the logarithmic bin-width axis. This pattern was repeated across V1-V3, and was not apparent in MT+/V5.

To quantify the effect of decreasing the number of bins used in the 'retinotopic' (polar angle) compared to 'random' averaging, we fitted an exponential growth function of the form

$$y(x) = p_c + (A - p_c)\left[1 - exp(-x/B)\right] \tag{4.2.1}$$

where y is the classification accuracy as a function of x (the number of bins), the parameters A and B are the asymptote and the exponential constant in units of number of bins, respectively, and $p_c = 0.125$ is the chance level. A function of this form was

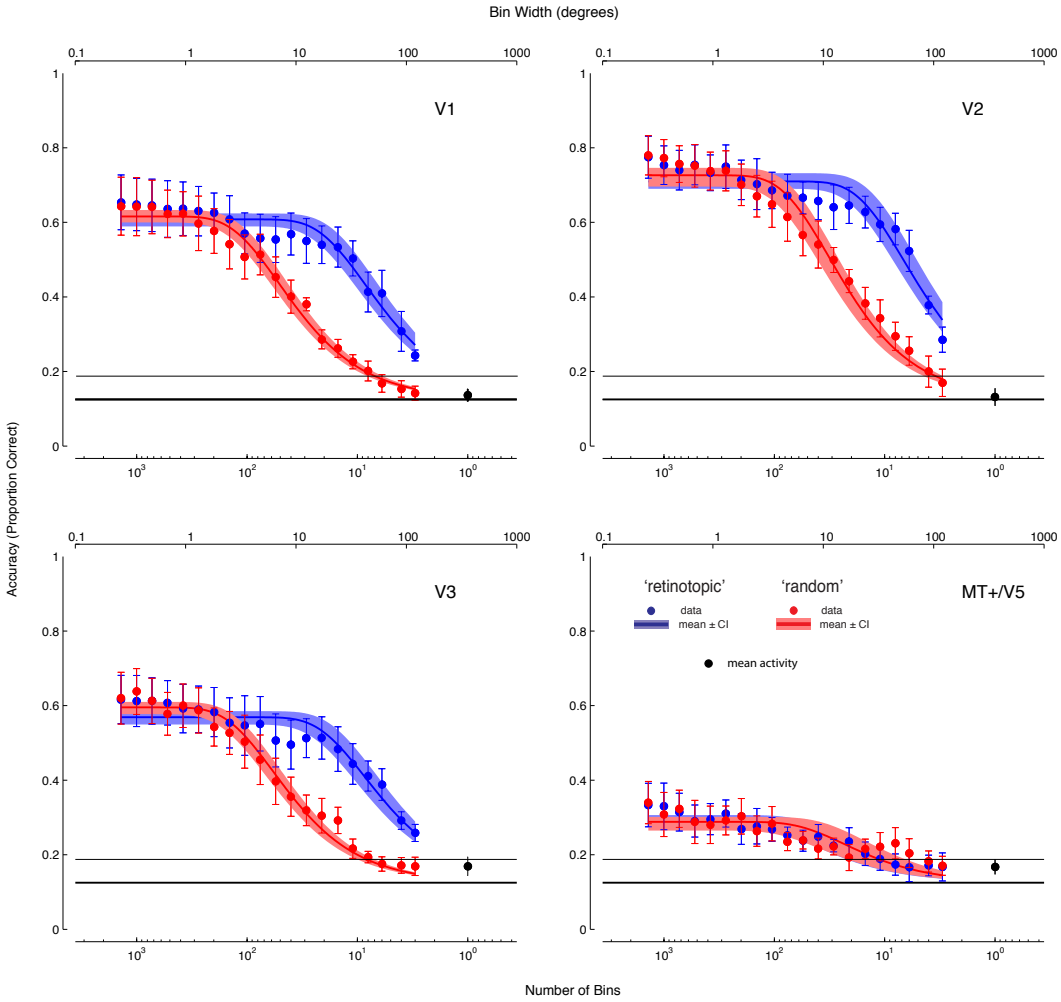

Figure 4.10: Effect of binning and averaging 500 voxels from V1-V3 and MT+/V5 based on their visual angle preference. Each panel shows data from a different ROI (V1, V2, V3, and MT+/V5). Blue symbols, binning and averaging by preferred visual field angle with increasing bin size (note log scale). Blue lines and shaded regions, exponential fit and 95% confidence intervals (see Methods for details), Red symbols, equivalent averaging but with shuffled preferred visual field angle labels, therefore providing a randomized control. Red lines and shaded regions, exponential fit and 95% confidence intervals for shuffled control.Black line, chance level proportion correct (0.125), grey line indicates 95th percentile of null distribution. Black symbols in each panel, control indicating that classification accuracy drops to chance level when using mean fMRI response across each ROI.

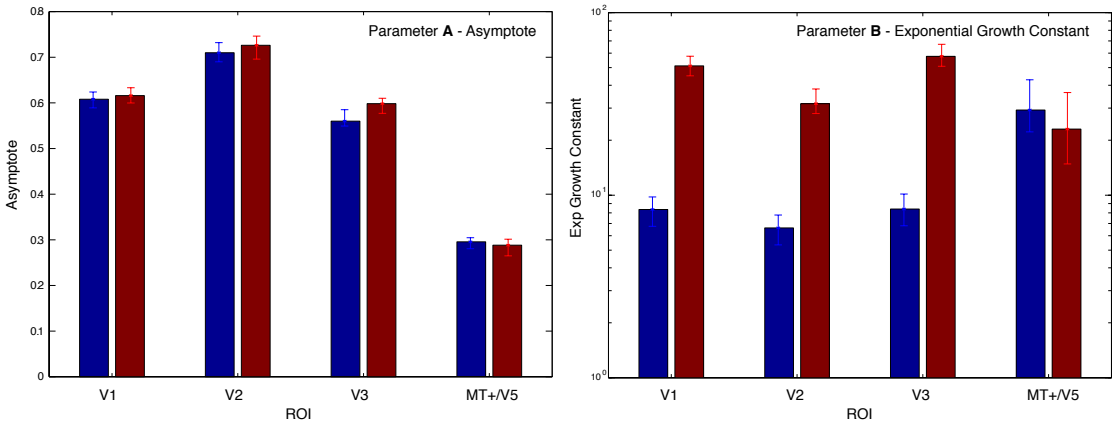

Figure 4.11: Mean parameters A (right panel) and B (left panel) from exponential fit to the data in Figure 4.10. Error bars represent the 95% Confidence Intervals from the bootstrapping procedure. Asymptotic classification accuracy (top panel) in the 'retinotopic' (blue bars) and 'random' (red bars) binning and averaging schemes is the same across all ROIs, while the scale parameter B (representing the number of bins required for successful classification) significantly differs (as judged by the 95% CIs from the bootstrapping procedure) between 'retinotopic' and 'random' binning for V1-V3 but not MT+.

chosen as it has been shown to fit well the pattern of results obtained by incrementing the number of voxels used in classification analysis (Mannion et al., 2009). We used a non-parametric bootstrapping technique to estimate 95% (BC_a, bias corrected and accelerated) confidence intervals for the fitting parameters A and B. Note that differences in the exponential constant between the two conditions appear as shifts on the log-scale plots used here. Also note that an increase in the number of bins (super-voxels) used corresponds to a decrease in the number of polar angles averaged together.

We computed 5,000 bootstrapped replications for each fitted function by re-sampling the residuals; the resulting distributions were used to generate a confidence interval for each parameter, without assumptions about the distribution of the raw data or the residuals (Efron and Tibshirani, 1993). The parameter A, representing the asymptote of each curve, did not significantly differ across the two averaging schemes, whereas the exponential growth parameter B showed a significant difference in the two averaging schemes in early visual areas (Figure 4.11). Note that the error bars in those plots represent 95% confidence intervals. We considered fitting parameters with non-overlapping confidence intervals as statistically significantly different.

4.2.3 Discussion

To assess whether a coarse-scale signal would be sufficient for successful classification, rather than a high spatial frequency signal such as columnar distribution, we repeated the classification after averaging voxels within an ROI based on their preferred visual field angle. This process removes any of the high-frequency information, such as the slight biases for different directions of motion that each voxel would obtain from the distribution of motion selective columns. Because our sampling resolution (1.5 mm isotropic) was larger than that needed to resolve the expected pattern of columnar organization, the spatial distribution of any small biases would appear random (Kriegeskorte et al., 2010). While averaging together responses from different voxels, small biases should cancel out while any information that is correlated with retinotopy (visual field angle preference) should be preserved. There are several possible mechanisms that could explain the presence of this coarse-scale information, including biases for cardinal (Schluppeck and Engel, 2003) and radial directions of motion (Clifford et al., 2009; Raemaekers et al., 2009). A bias for cardinal motion directions is unlikely: when the mean amplitude of the ROI was used for classification, classification accuracy was at chance. To check for a radial bias, we examined the largest weight that each voxel or super-voxel contributed to each direction detector, and observed how these weights varied depending on the polar angle of visual field the voxel or super-voxel represented, based on its retinotopic phase or the range of phases that super-voxels bin encompassed. When the voxel weights were displayed on the cortical surface, no pattern could be observed, but as the voxels were binned over increasing ranges of polar angle to form the super-voxels, biases for directions of motion radial to the supervoxel bin's mean polar angle began to emerge (Figure 4.12). This is consistent with the previous findings that any radial biases are quite small, and only directly observable using a combination of methods and sufficient amounts of data (Freeman et al., 2011; Raemaekers et al., 2009), or when the data have been blurred considerably (Swisher et al., 2010).

We found that averaging voxels together based on their visual field preference preserved classification performance compared to random averaging of the equivalent number of voxels. By fitting curves to the data, we were able to calculate 2 parameters for each ROI/binning method combination: a parameter A representing the asymptote and a parameter B representing the exponential growth constant, a change in which represents a shift of the curve along the log-scale x axis in Figure 4.10. The lower the B value, the lower the number of bins required for a given level of classification (which corresponds to a greater number of voxels averaged together). The B parameter was

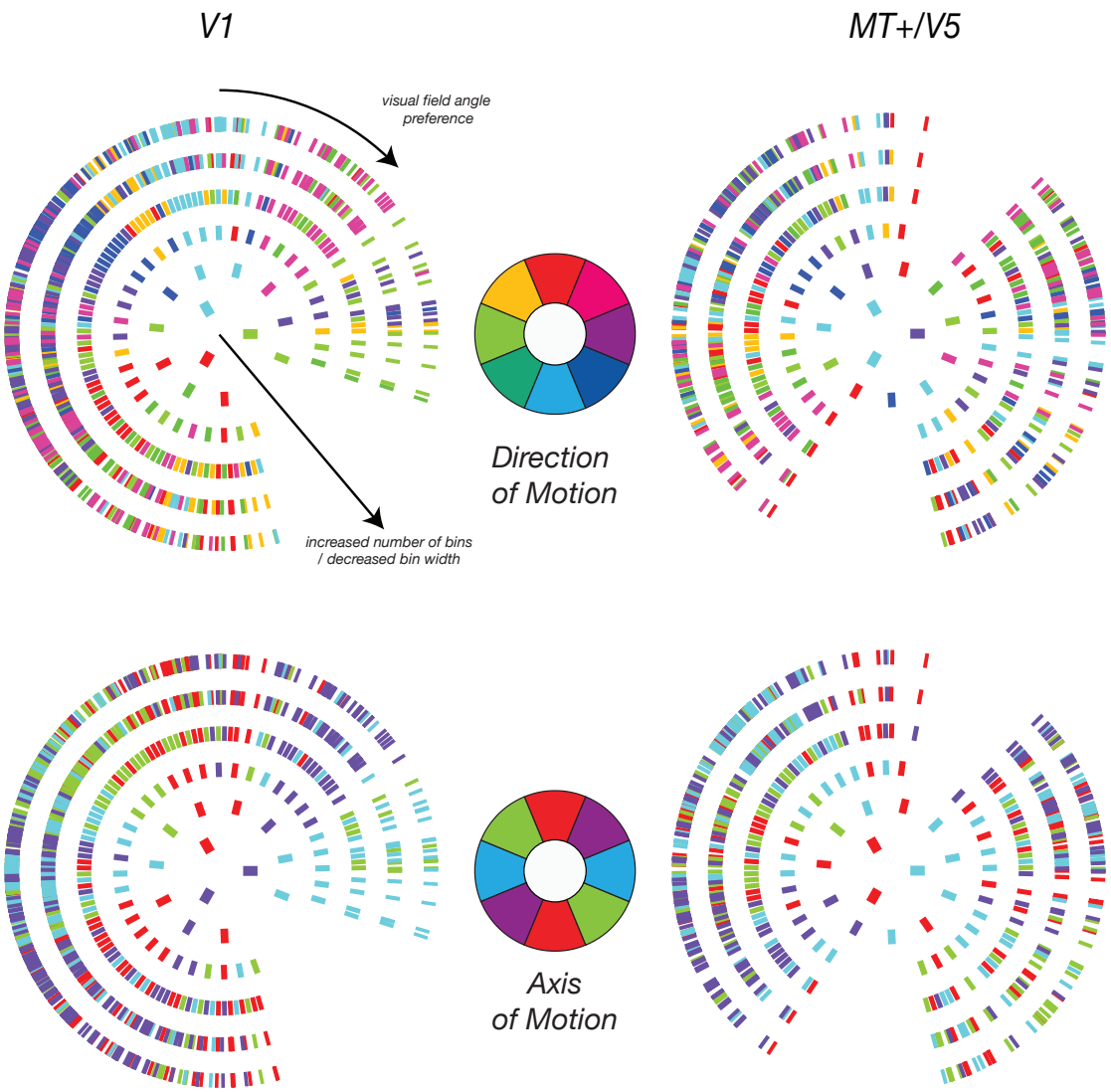

Figure 4.12: Largest weights (colors) for 'direction of motion' or 'axis of motion' detectors as a function of visual field preference for (super-)voxels. Left column shows data for V1; right column for MT+. Top, analysis considering 'direction of motion' (0-360 degrees); bottom, 'axis of motion' (0-180 degrees). Polar plots show weights for (super-)voxels according to their phase value from the retinotopic localizer. Angle (theta) for each symbol, phase value from the retinotopic localizer (mean across component voxels making up super-voxels). Eccentricity of ring (R), different levels of binning: from outer ring (no binning) towards the centre of the plot (binning, large phase bins). Colour coding shows which direction detector that (super-voxel) voxel has the largest weighting. As the bin widths increase (towards centre), a bias for radial directions of motion becomes apparent. The analysis for axis of motion reveals a stronger radial bias (cf. bottom panels versus top panels).

significantly lower (as assessed by bootstrapping) for the retinotopic-binning method compared the the random-binning method in V1-V3, indicating that a smaller number of bins were required for successful classification in this condition. Averaging voxels together in a way that preserves the signal at the scale of the retinotopic map allows significantly improved performance compared to averaging together the equivalent number of voxels at random, indicating that there is information at this spatial scale sufficient for successful classification.

Our results are comparable to the findings of Freeman et al. (2011), who found almost unaffected performance even when bin width was increased to ~60° of visual field angle (see their Figure 5a). However, for both our averaging schemes, 'retinotopic' and 'random', classification performance dropped significantly for the largest bin sizes. There are several reasons for this difference. Firstly, we used blocks of motion-defined stimuli rather than slowly rotating gratings and secondly, the phase values we used for binning voxels for preferred visual field angle were based on 2 within-session scans, rather than a separate full-session of retinotopy, which may have added variability to our voxelwise visual field measurements.

A slight bias for cardinal directions of motion has been reported by Schluppeck and Engel (2003), although the finding of chance performance when the average time-series for each ROI was used for classification suggests that this is not the basis for the classification results.

Recent studies have suggested biases for radial directions of motion in human visual cortex (Clifford et al., 2009; Raemaekers et al., 2009; Sun et al., 2012). Given that Freeman et al. (2011) demonstrated a radial orientation bias was both *necessary* and *sufficient* for classification of orientation in V1, an analogous radial bias for direction of motion is a likely candidate signal underlying the motion classification. Another possibility is that orientation signals in the form of motion 'streaks' may be the signal driving classification of non-opponent motion (Apthorp et al., 2013). If this were the case, a radial bias for orientations could be the mechanism underlying both orientation and (non-opposing) motion classification.

Raemaekers et al. (2009) reported a radial bias in early visual cortex, but not for MT+/V5. In our data, there is no strong evidence of an advantage for retinotopic averaging over random averaging in MT+/V5 suggesting a weakened or absent contribution from radial bias in driving classification performance in this region. However, there are some limitations to this interpretation. Comparing results across different visual cortical areas is problematic. While MT+/V5 is generally more selective for direction of motion than early areas such as V1 (based on the proportion of direction-selective cells found

in macaque electro-physiology (Dubner and Zeki, 1971; Zeki, 1974)), differences in the columnar organization or the amplitude of the BOLD response may lead to lower classification accuracies. The ability to decode stimulus properties from signals in a given area may (but not need) indicate selectivity to those properties in a given area. Conversely, however, a failure of decoding cannot be taken as evidence for a lack of selectivity (Bartels et al., 2008).

4.3 Controlling for Eye Movements

One possible mechanism for inducing a retinotopic signal that could lead to artefactual classification is a systematic relationship between different stimulus directions and eye movements (Kamitani and Tong, 2006), where eye movements could lead to global shifts of activated regions in retinotopically organised brain areas. Freeman et al. (2011) measured eye movements during their orientation classification experiments, and found that their results could not be explained through eye movements. Although in the current experiment subjects performed a demanding contrast discrimination task at fixation during scanning, their eye movements were not monitored due to technical limitations of our high-field MRI setup. We could therefore not be sure that subjects had not moved their eyes during stimulus presentations. To control for this possible artefact, in a separate experiment, we tested (a) whether the ability to classify direction of motion of stimuli could be due to a specific pattern of eye movements and (b) whether the pattern of results was particular to our choice of full-field stimuli.

4.3.1 Methods

Participants, Stimuli & Procedure

Two of the subjects used in the motion classification experiment were also used for the split-hemifield experiment. Stimuli were identical to the block-classification stimuli, except that the visual field was split into two apertures along the vertical meridian, with the dots in each aperture drifting in different directions, independently chosen, for each 16s block. The stimuli used in the control experiment had the same properties as in the main experiment, with the modification that the display was split into left and right hemifields along the vertical meridian, with the dots in each hemifield aperture drifting in different directions, independently chosen, for each 16s block. This was done to exclude the possibility that eye movements (which could not be measured in the 7 T scanner) could explain the classification results; in these stimuli no singular pattern of eye movements could lead to reliable decoding of direction (Swisher et al., 2010).

4.3.2 Results and Discussion

ROIs corresponding to the left and right visual fields were analysed and binned separately over the polar angle ranges represented in each ROI. We found the same pattern of results for these stimuli. Firstly, the direction of motion could be decoded from the

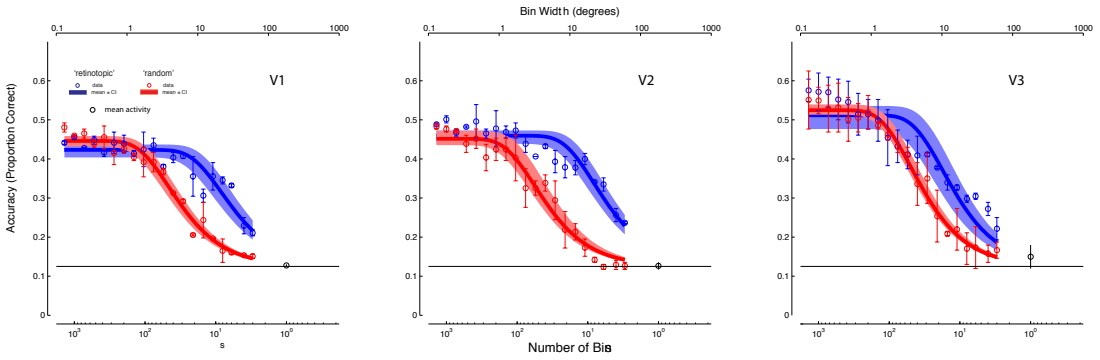

Figure 4.13: Control experiment using split-hemifield stimuli. Same conventions as in Figure 4.10. Graph shows the effect of binning and averaging 500 voxels from V1-V3 based on their visual field preference. Blue symbols, binning and averaging by preferred visual field angle with increasing bin size (note log scale). Red symbols, equivalent averaging but with shuffled preferred visual field angle labels, therefore providing a randomized control. Lines, fit to data using Equation 4.2.1.

patterns of activity in the visual cortex as is evident by the asymptotic classification accuracy in Figure 4.13, which was significantly above chance level. Secondly, binning and averaging voxels according to their visual field preference showed preserved performance compared to averaging after random binning, as is evident from the rightward shift of the blue curve in the log-plot in Figure 4.13.

Binning and averaging voxels by visual field preference (as in the main experiment) still showed a benefit compared to random averaging (see data for V1-V3 in Figure 4.13). Therefore, the effects we describe here are unlikely to be due to any information afforded by eye movements correlated to the stimuli. Although the results for V1 and V2 still show an advantage for retinotopic averaging compared to random averaging, V3 shows a much reduced advantage. Whilst this may suggest that classification in V3 is in fact based on eye movements, upon examining the stack placement for the 2 subjects, it was found that the dorsal side of V3 was not fully covered, which would explain the reduced overall performance in this area.

4.3.3 Discussion

One possible explanation for a bias for radial directions contributing to classification of motion directions could be the influence of eye-movements systematically altering the activated area of retinotopic cortex. To investigate this, the classification experiment was repeated using stimuli that should not allow a consistent mapping of eye

movements to direction of motion, by displaying different directions of motion in each visual hemifield (Swisher et al., 2010). Even with this control, elevated classification accuracies were found when voxels in each ROI were averaged together based on their retinotopic phase, compared to when voxels were averaged randomly (Figure 4.13). This indicates that both neither successful classification, nor the preservation of performance after retinotopic averaging, are due to systematic eye movements. The results therefore suggest that a coarser scale signal, organised on the basis of retinotopic polar angle, is in fact a major contributor to successful classification of the translation direction of drift for dot patterns.

4.4 General Discussion

We replicated the classification of motion-direction experiment by Kamitani and Tong (2006) at a higher field strength (7T as opposed to 3T). The main rationale was to increase the resolution from 3 mm isotropic to 1.5 mm isotropic, representing an 8-fold decrease in voxel volume. We found that a resolution of 3 mm versus 1.5 mm isotropic showed very little difference in performance (Figure 4.8), indicating that resolution greater than that typically offered by standard strength MRI does not offer any clear benefit for multivariate classification. Although the link between voxel size and performance is not always straightforward (Kamitani and Sawahata, 2010; Kriegeskorte et al., 2010), this failure to see improvement at finer resolution could indicate that the information used by the classifier is not at a fine-scale. Low-pass filtering of our data using Gaussian blurring did show some decrease in performance, although performance was still above-chance when the largest filter kernel was used. Drawing conclusions about the scale of the signal used from classification performance after Gaussian blurring can be difficult due to a) the fact that Gaussian blurring may not fully isolate the desired frequencies (hence the well above chance performance even with the finest high-pass filtering) (Kamitani and Sawahata, 2010; Swisher et al., 2010) b) the potential variability in the effects of filtering due to head motion or other factors (Misaki et al., 2012) and c) the fact that the scale of a signal and the frequencies that it is represented at in an image may be different from that expected (Freeman et al., 2011; Kriegeskorte et al., 2010). Freeman et al. (2011) showed that low- and high-pass filtering had similar effects on classification of orientation and angular position, the latter of which is known to be represented at a coarse-scale on the cortical surface. They hypothesized that even a coarse-scale organisation would be represented in part at finer spatial frequencies and hence susceptible to low-pass filtering.

Freeman et al. (2011) suggested that classification for certain orientation stimuli maybe be based on a coarse-scale bias for radial orientations, rather than fine-scale information based on the distribution of orientation selective columns. Given that evidence for a bias for radial directions of motion has also been demonstrated (Raemaekers et al., 2009), we performed a further analysis to ascertain whether direction of motion classification could be explained by a radial bias. Averaging voxels together based on shared polar angle representation did how some reduction in performance, but significantly less than when compared to an equivalent level of averaging that did not preserve polar angle. This result demonstrates the sufficiency of a coarse-scale, retinotopically organised signal (for example a bias for radial directions of motion) for motion classification.

One potential mechanism to induce a coarse-scale signal is a mapping of eye-movements with certain directions of motion, by changing the area of retinotopic cortex that is activated in response to each direction of motion (Kamitani and Tong, 2006). To control for this, we repeated the experiment using stimuli where different directions of motion were displayed in each visual hemifield, meaning no single pattern of eye-movements could be associated with each direction in each hemifield. We still found that a coarse retinotopic signal was sufficient to explain motion classification in V1 and V2, suggesting that eye-movements were not the cause of the our earlier result.

The experiments in this chapter add to recent evidence that multivariate classification results for certain stimuli may be based on coarse-scale signals rather than on signals deriving from the distribution of stimulus-selective neurons or columns (Gardner, 2010; Op de Beeck, 2010). While results from low- and high-pass filtering experiments can be equivocal with regard to the scale of the signal (Freeman et al., 2011; Kamitani and Sawahata, 2010; Misaki et al., 2012; Swisher et al., 2010), experiments directly studying a coarse-scale signal (Freeman et al., 2011) and simulations (Chaimow et al., 2011) indicate that such coarse-scale signals are both necessary and sufficient for successful classification, especially at the levels seen in most classification studies.

Chapter 5

Classification of non-translational motion

5.1 Classification of rotation and 'spiral' motion

The experiments in Chapter 4 demonstrate that coarse-scale retinotopically arranged biases may explain some or all classification results for translational motion. However, classification is also possible using a wide range of different stimuli, including those that lack any component that would be picked up via a radial bias. For example, using stimuli such as Glass patterns, where each dot pair is oriented relative to radial angle (see Fig 5.1), clockwise and anti-clockwise patterns can be discriminated using a multivariate classifier (Mannion et al., 2009). Successful classification of rotating clouds of dots, which similarly would have balanced radial components, has also been demonstrated (Kamitani and Tong, 2006; Seymour et al., 2009). Additionally, an experiment using radially modulated luminance noise (Hong et al., 2011) was able to show successful motion classification.

Results such as these have been used as evidence that coarse scale biases are not necessary for successful classification, as stimuli that have balanced radial components are still able to be classified using multivariate techniques (Clifford et al., published online 26 April 2011). However, it has been argued that stimuli typically considered to be radially balanced may still induce large-scale biases in visual cortex, co-varying with the retinotopic map, that could be used by a classifier (Merriam et al., 2012). If such large scale biases exist for stimuli that are ostensibly radially balanced, this provides further evidence that successful multivariate classification does not necessarily tap into fine-scale neural information.

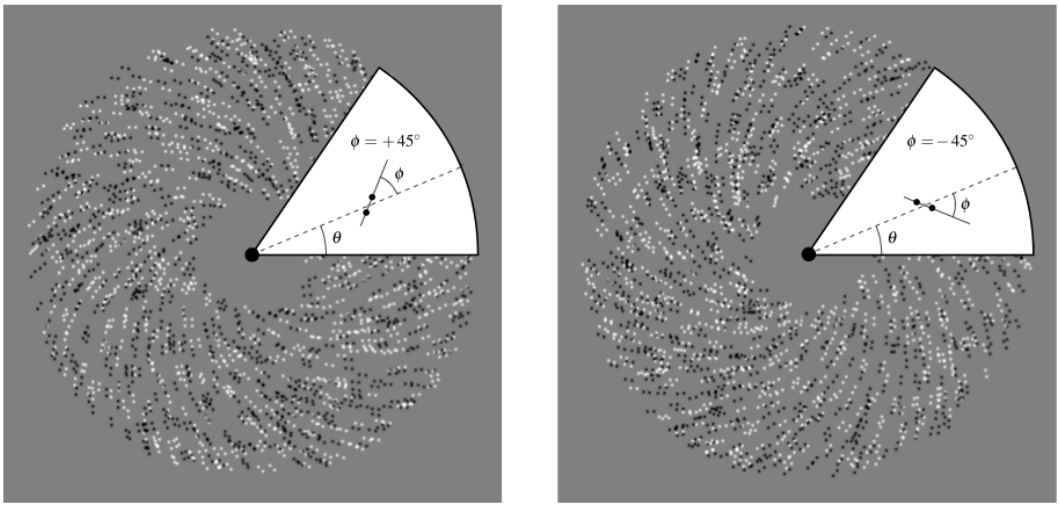

Figure 5.1: Example of radially balanced Glass pattern stimuli. Each dot pair is oriented at an angle (ϕ) of either +45 (left) or -45 (right), relative to the radial angle (θ), creating either anti-clockwise or clockwise glass patterns

We attempted to replicate the classification of radially balanced motion stimuli to investigate how classification of this kind would be affected by the retinotopic averaging regime used on the translational motion. If the classification of such stimuli is in fact unaffected by retinotopically arranged biases, then both the random and retinotopic averaging methods should have the same effect on performance. If however classification performance is reliant on a signal at the same scale as, but not actually based on, the radial bias then the retinotopic averaging may afford some benefit. In the initial experiment we tested classification for rotational motion stimuli similar to that used by Kamitani and Tong (2006), and a motion version of the glass pattern stimuli used by Mannion et al. (2009). We then followed this up with an experiment using stimuli that could be classified along a number of dimensions including sign of rotation, and orientation of the motion trajectory relative to radial angle.

5.1.1 Methods

Participants, Stimuli & Procedure

Two experienced subjects took part in the initial experiment. Stimuli were identical to that used in the translational motion experiment, except in the way that dot motion was defined. On each frame, the angle of the step that each would make was defined relative to the radial angle of that dot, in a similar fashion as for the Glass patterns in Mannion et al. (2009). A (ϕ) of 0 would lead to expanding dots, a (ϕ) of

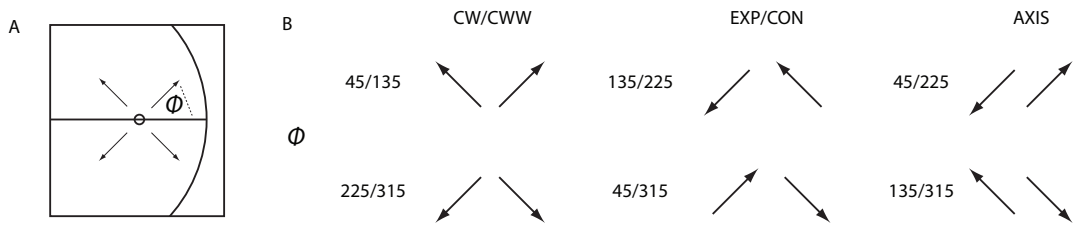

Figure 5.2: Demonstration of the stimuli used in the second experiment and the different dimensions the classification could be performed on. A: A dot in the visual field had a motion trajectory defined by an angle (ϕ) relative to its polar angle in the visual field. The ϕ used in the experiment corresponded to clockwise/anticlockwise expansion/contraction, or 'spiral motion'. B: By Combining the stimuli across various dimensions, different classifications along different dimensions could be performed.

$\pm$90 would yield anti-clockwise/clockwise rotation, and a (ϕ) of $\pm$45 would yield anti-clockwise/clockwise 'spiralling' motion that combined expansion and rotation. Participants viewed two examples of each stimulus per scan. Participants performed the same fixation task as in the original experiment.

In a follow-up experiment, two subjects (one who took part in the initial experiment and one naive subject) viewed stimuli with (ϕ) of 45, 135, 225 and 315 degrees relative to polar angle, corresponding to 'anticlockwise expansion', 'anticlockwise contraction', 'clockwise contraction' and 'anticlockwise expansion'. These stimuli could either be classified in a 4-way classification or in a number of 2-way classifications along a number of different dimensions (Figure 5.2).

Data Analysis

Patterns for classification were formed in the same way as for the original experiment (Chapter 4), by taking the temporal mean of the 8 TRs following each stimulus presentation shifted by 2 TRs to account for the hemodynamic lag. The patterns from each scan were z-scored on a per scan basis for each voxel. Leave-one-run-out cross validation was used with a linear SVM to assess the classification accuracy. The retinotopic binning analysis was also repeated for the radially-balanced motion classification.

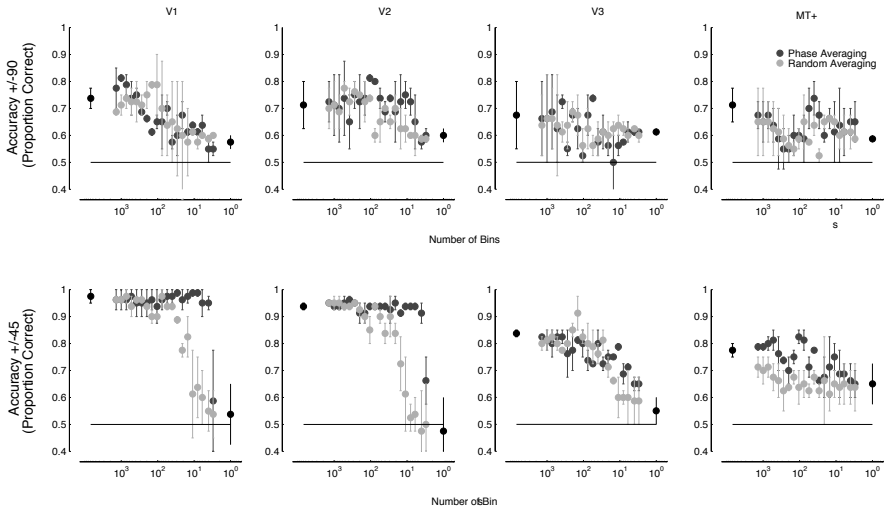

Figure 5.3: Results from the initial radially balanced motion experiment. Top Row: Comparing $-90°$ (anticlockwise) with $90°$ (clockwise) rotation. Dark grey markers show classification after retinotopic averaging, light grey markers show random averaging. Bottom Row: $-45°$ (anticlockwise) versus $45°$ (clockwise) 'spiral' motion. Error bars represent the standard error of the mean across subjects (n=2).

5.1.2 Results

We first compared $-90°$ (anticlockwise) with $90°$ (clockwise) rotation (Table 5.1, Top Row). Above chance classification as determined by a permutation test was found for both subjects in V1 and V2, although significant classification was only found one subject in V3 and MT. Above chance classification was found in all ROIs for the $\pm 45°$ condition (Bottom Row), with accuracies for this discrimination being higher than the $\pm 90°$ condition in earlier areas. This is consistent with the findings of Kamitani and Tong (2006) who showed that classification of orthogonal motion is more accurate than opposite direction classification.

+/- 90	V1	V2	V3	MT
ab	**0.775 (<0.001)**	**0.800 (<0.001)**	**0.800 (<0.001)**	**0.775 (<0.001)**
jb	**0.700 (0.002)**	**0.625 (0.046)**	0.550 (0.258)	0.650 (0.051)

+/- 45	V1	V2	V3	MT
ab	**1.000 (<0.001)**	**0.925 (<0.001)**	**0.850 (<0.001)**	**0.750 (0.001)**
jb	**0.950 (<0.001)**	**0.950 (<0.001)**	**0.825 (<0.001)**	**0.800 (0.001)**

Table 5.1: Mean accuracies and p values (in brackets) for each ROI and subject in the spiral classification experiment.

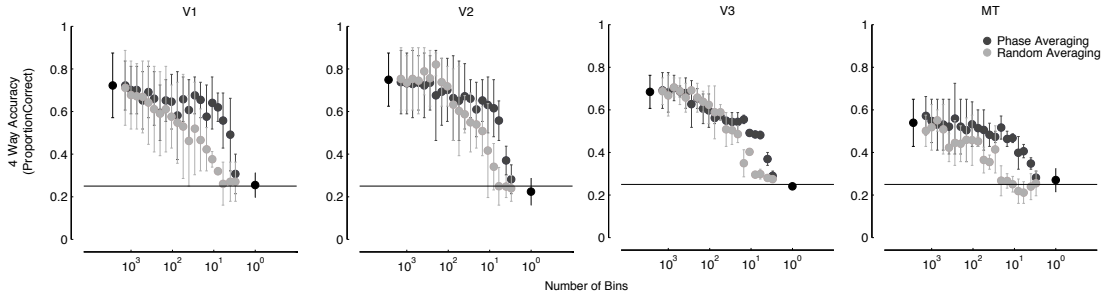

Figure 5.4: Results from the 4 way classification analysis of the spiral motion stimuli. Dark grey markers show retinotopic averaging, light grey markers show random averaging. Error bars represent the standard error of the mean across subjects (n=2).

When the retinotopic averaging analysis was applied (Figure 5.3), no consistent preservation of performance was found in the retinotopic averaging condition compared to the random averaging condition for the $\pm 90^\circ$ condition (Top Row). This is consistent with rotation stimuli being balanced in their radial components, therefore meaning that no retinotopically organised bias for radial directions of motion should be able to contribute to the classification (Kamitani and Tong, 2006).

We then compared this with classification for spiral motion with a narrower angular separation, similar to the glass patterns used in the orientation case (Mannion et al., 2009). $\pm 45^\circ$ classification also showed preserved performance for retinotopic averaging in early visual areas, even though stimuli were radially balanced (Bottom Row).

A 4 way classification experiment was then undertaken using one repeated subject and one novel subject, with polar angle relative directions equally spaced (45°, 135°, 225°, 315°) (Figure 5.2). A simple 4 way classification showed above chance classification in all ROIs for each subject (Table 5.2), and a benefit for retinotopic averaging compared to random averaging (Figure 5.4).

4 way	V1	V2	V3	MT
ab	**0.875 (<0.001)**	**0.875 (<0.001)**	**0.762 (<0.001)**	**0.650 (<0.001)**
ms	**0.554 (<0.001)**	**0.625 (<0.001)**	**0.607 (<0.001)**	**0.393 (0.012)**

Table 5.2: Mean accuracies and p values (in brackets) for each ROI and subject in the 4-way spiral classification experiment.

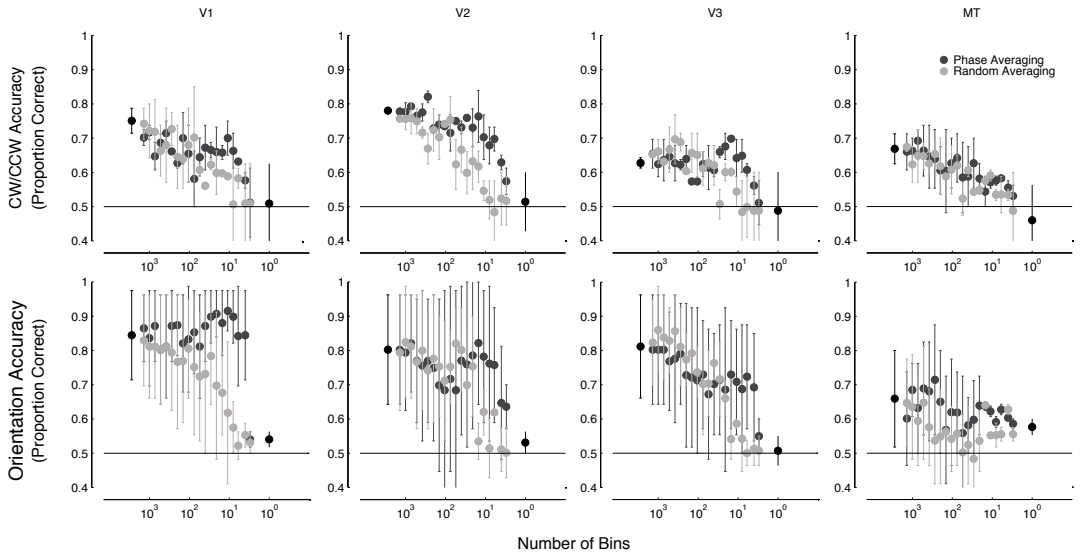

Figure 5.5: Results from performing the classification across 2 different stimulus dimensions in the 4-way spiral classification experiment: CW/CCW orientation (Top Row), and Axis of Motion Orientation (Bottom Row). Dark grey markers represent retinotopic averaging, light grey markers represent random averaging. Error bars represent the standard error of the mean across subjects (n=2).

By combining the stimuli to be discriminated in different configurations (Figure 5.2), the effect of retinotopic binning on discrimination of direction of rotation and the orientation of motion-trajectory could be determined. Table 5.3 shows the classification results and significance values from a permutation test for each subject and ROI for the rotation (Top Row) and Orientation (Bottom Row) conditions. Both subjects showed above chance performance in V1 and V2, with some variability in V3 and MT.

orientation	V1	V2	V3	MT
ab	**0.975 (<0.001)**	**0.963 (<0.001)**	**0.963 (<0.001)**	**0.800 (<0.001)**
ms	**0.714 (<0.001)**	**0.643 (0.005)**	**0.661 (0.003)**	0.518 (0.386)

cw/ccw	V1	V2	V3	MT
ab	**0.787 (<0.001)**	**0.775 (<0.001)**	0.613 (0.068)	**0.713 (0.002)**
ms	**0.714 (<0.001)**	**0.786 (<0.001)**	**0.643 (0.001)**	**0.625 (0.019)**

Table 5.3: Mean accuracies and p values (in brackets) for each ROI and subject in the 4-way spiral classification experiment across the 2 possible classification dimensions.

The binning procedure from the previous experiments were repeated for all the different stimulus configurations (Figure 5.5). Classification for direction of rotation (Top Row) again showed no advantage for retinotopic averaging, whereas classification of

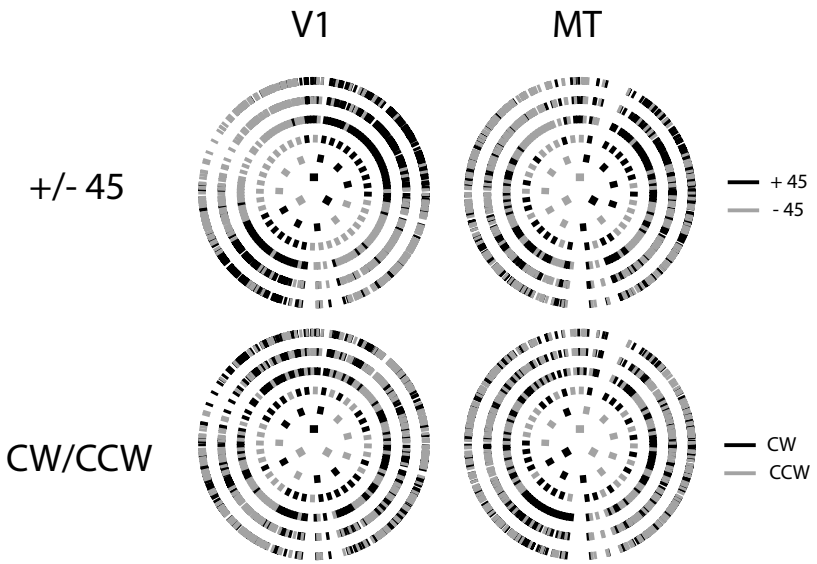

Figure 5.6: Largest weights for $\pm 45^{\circ}$ or 'direction of rotation' detectors as a function of visual field preference for (super-)voxels. Left column shows data for V1; right column for MT+. Top, analysis considering $\pm 45^{\circ}$; bottom, 'sign of rotation' (cw/ccw). Same plotting format as for Figure 4.12. The analysis for $\pm 45^{\circ}$ reveals a bias for a given stimulus depending on visual quadrant.

the axis of motion relative to radial angle (Bottom Row) did show an advantage for retinotopic averaging in V1.

5.1.3 Discussion

Classification of stimuli without radial biases is possible, congruent with previous results using orientation (Mannion et al., 2009) and motion (Hong et al., 2011; Kamitani and Tong, 2006; Seymour et al., 2009) stimuli. Classification of 'simple' rotation (Figure 5.3, Figure 5.5) is possible in all areas tested, similar to previous studies. In addition, the angle of the motion relative to the polar angle of the dots (ϕ) can also be successfully classified (Figure 5.3, Figure 5.5), mirroring results seen using oriented Glass patterns (Mannion et al., 2009). Stimuli such as these are assumed to be balanced in their radial components, and therefore should not be classifiable on the basis of a coarse scale signal such as a radial bias in visual cortex (Clifford et al., published on-line 26 April 2011). However, there have been suggestions that even stimuli that are unbiased in terms of their radial components may not be completely free of large-scale biases that can still contribute to successful classification.

To investigate whether an organised coarse-scale signal contributes to the classifica-

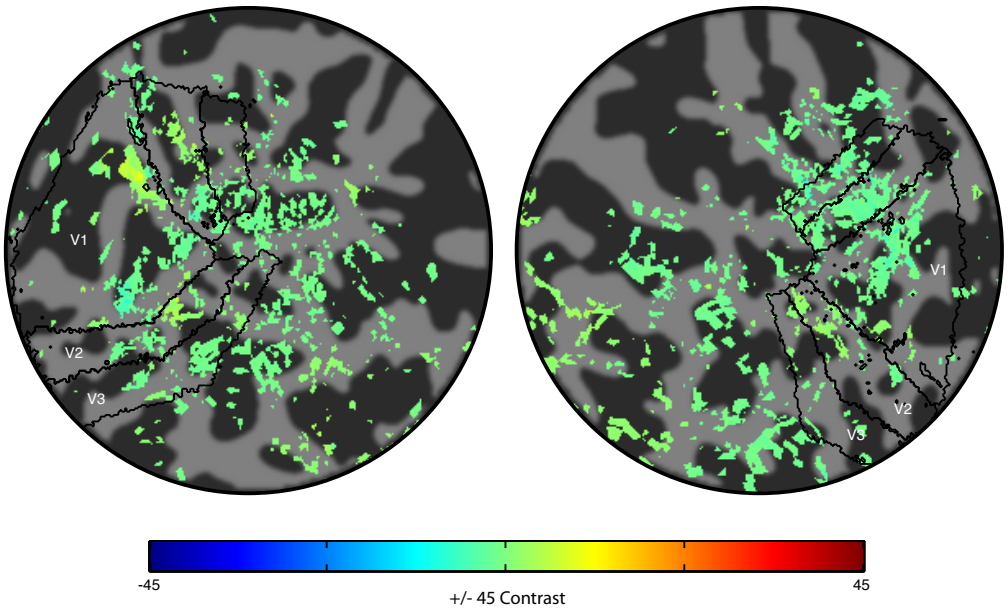

Figure 5.7: Voxelwise univariate contrast of ±45° displayed on the cortical surface. Values are thresholded according to an F-Test ($p < .05$, no correction for multiple comparisons).

tion of non-translational motion such as rotation, we performed the same retinotopic averaging analysis we performed in Chapter 4 (Freeman et al., 2011), averaging voxels together on the basis of their retinotopic phase or at random prior to classification. The results from the initial experiment showed no benefit for retinotopic averaging compared to random averaging for ±90 classification (Figure 5.3), indicating no coarse-scale signal being used. However, the same analysis done for ±45 classification did indicate a benefit for retinotopic averaging in V1 and V2, indicating a coarse-scale signal co-varying with polar angle. This result is consistent with the findings of Merriam et al. (2012) for orientation, who found large-scale organization of preference for different spirals in some subjects, with the upper and lower portions of retinotopic V1 showing opposite preferences. Examining the relation between the weights used by the classifier and the polar angle phase value of the voxels/supervoxels the weights arise from (Figure 5.6), this retinotopic bias can be seen for the ±45 stimuli, with a preference for opposite stimuli in alternating quadrants in V1. Such a bias is not apparent in MT+, or for the classification of rotational stimuli.

Figure 5.7 shows the ±45° univariate contrast weights for a single subject displayed on the cortical surface, showing a potential retinotopically organised signal for classification of this type.

For the follow-up experiment using 4 equally spaced angles of rotation, retinotopic

averaging showed a benefit for all visual areas, including a slight benefit in MT (Figure 5.4). Comparing the effect of averaging on the different classification dimensions showed benefits for classifying the axis of the motion (45/225° versus 135/315°) in V1, with slight benefits apparent in V2 and V3, and no benefit apparent in MT+. No clear benefits for retinotopic averaging were seen when classifying the sign of rotation (45/90° versus 225/315°) in any visual areas.

These results indicate that classifying the angle of rotation, as opposed to the sign of motion along that axis, shows a benefit for retinotopic averaging when compared to the equivalent amount of averaging done randomly. This indicates that a coarse-scale signal exists that co-varies with the retinotopic polar angle map and is sufficient for successful classification (Freeman et al., 2011). This is consistent with the finding by Merriam et al. (2012) that a coarse scale bias for clockwise versus anticlockwise glass patterns existed in V1 and could account for the successful classification of clockwise and anticlockwise glass patterns demonstrated by Mannion et al. (2009). Whether the evidence for a coarse-scale bias represents a motion-specific signal, or is based on the effect of orientation-artefacts from 'motion-streaks' (Apthorp et al., 2013) is not clear from these results.

We found no consistent benefit for retinotopic averaging when the classifications involved the sign of motion rather than the angle, i.e for simple rotation. This could indicate that no coarse-scale signal exists for classification of this type, meaning that classification of stimuli of this type (Hong et al., 2011; Kamitani and Tong, 2006; Seymour et al., 2009) did in fact rely on signals unrelated to any coarse-scale information. Therefore it may be assumed that using stimuli of this type to investigate motion selectivity with MVPA methods is appropriate. Conversely, it may simply be that a coarse-scale signal exists, but could not be captured by averaging over polar angle. Performing the analysis over a different dimension (e.g. eccentricity) could potentially reveal a coarse-scale signal that can be used for classifying stimuli of this type.

Consistent with previous work, we found that non-translational motion could be classified using multivariate methods, even when such stimuli had balanced radial components. It has been argued that the classification of such stimuli indicates that radial biases are not necessary for classification (Clifford et al., published on-line 26 April 2011), contrary to findings for orientation (Freeman et al., 2011). However, we showed that for classification where the angle (ϕ) of the motion had to be classified and not the sign of the motion, a coarse-scale signal that persisted when voxels were averaged together on the basis of their retinotopic phase was sufficient for successful classification, consistent with recent findings for orientation (Merriam et al., 2012). Plotting the

largest weights from the classifiers as a function of polar angle revealed the potential form of such a bias for these stimuli (Figure 5.6, Top Row).These results indicate that even when using stimuli that should be balanced in their radial components, coarse-scale biases sufficient for classification may still exist. This relates to a broader issue that must be taken into consideration with multivariate classification, that successful classification is not definitively based on signals arising from fine-scale neural architecture, but may in fact be based on signals arising at a much coarser scale, or at a variety of scales. Even in cases where coarse-scale biases are believed to have been controlled for, care should be taken when interpreting classification results.

There was no indication for the use of a coarse scale signal for the classification of stimuli that differed in the sign of their rotation (Figure 5.6, Bottom Row). Such stimuli have successfully been used to show the effects of attention on the classification of ambiguous stimuli (Kamitani and Tong, 2006), which visual areas can encode the conjunction of colour and motion for a given stimulus (Seymour et al., 2009), and which areas demonstrate generalisation of classification between first and second order motion (Hong et al., 2011). These stimuli can thus be appropriately used in the study of a number of experimental questions, and could potentially ensure no contribution from coarse scale signals. However, they may not be appropriate to study certain aspects of visual processing, i.e. the tuning of the motion detectors (Kamitani and Tong, 2006).

Chapter 6

Classification of Pattern Motion

A network of areas in the brain are involved in the computations underlying our perception of object motion, as demonstrated by a range of electrophysiological (Dubner and Zeki, 1971; Hubel and Wiesel, 1962, 1965) and neuroimaging (Braddick et al., 2001; Kamitani and Tong, 2006) data. This leads to the question of the nature of the processing in each visual area, and particularly how the fairly complex motion of objects can be computed from the fairly simple inputs to the visual system, specifically the direction-selective neurons in V1. Hubel and Wiesel (1962, 1965) identified direction selective neurons in early visual cortex using drifting oriented bars of light. Given the nature of neurons in V1, particularly the restricted size of their receptive fields, they should be unable to accurately signal the direction of motion of an object due to what is known as the aperture problem (Marr and Ullman, 1981).

The direction of motion of a moving edge when viewed through a small aperture depends critically on the orientation of the edge relative to the aperture, because moving edges seen through a small aperture appear to be moving orthogonally to their orientation (Figure 6.1). This phenomenon is known as the aperture problem (Marr and Ullman, 1981). The restricted, oriented receptive fields in V1 and other early visual areas suffer from the aperture problem, responding only to the motion energy at each neuron's preferred orientation. Therefore to recover the true motion of an object whose extent goes beyond the size of the individual RFs, the aperture problem must be solved. One solution is to combine the local velocity estimates into a single estimate of object velocity, and several studies have identified area MT/V5 as a possible locus for such a computation.

Initial studies of possible solutions to the aperture problem used a set of stimuli called plaids (Figure 6.2), formed from overlaid sinusoidal gratings with different orientations

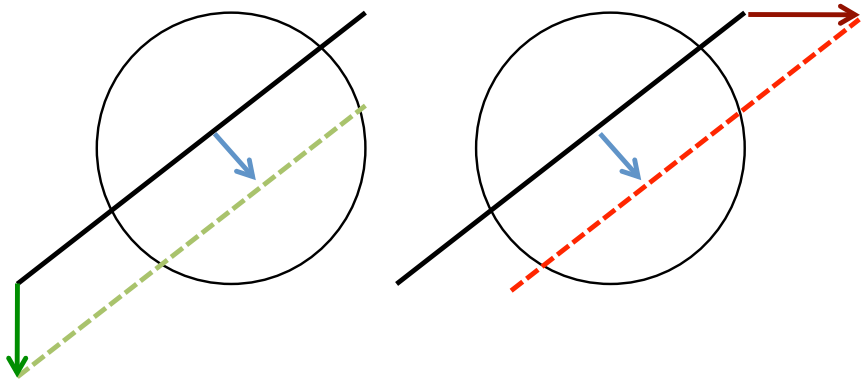

Figure 6.1: Schematic demonstrating the aperture problem. The circle represents a direction selective cell with a limited RF (e.g a V1 cell), with an oriented line drifting through it. The motions on the left and right (green and red lines respectively) both lead to the same direction being signalled by the cell (blue arrows), due to the restricted nature of the RF.

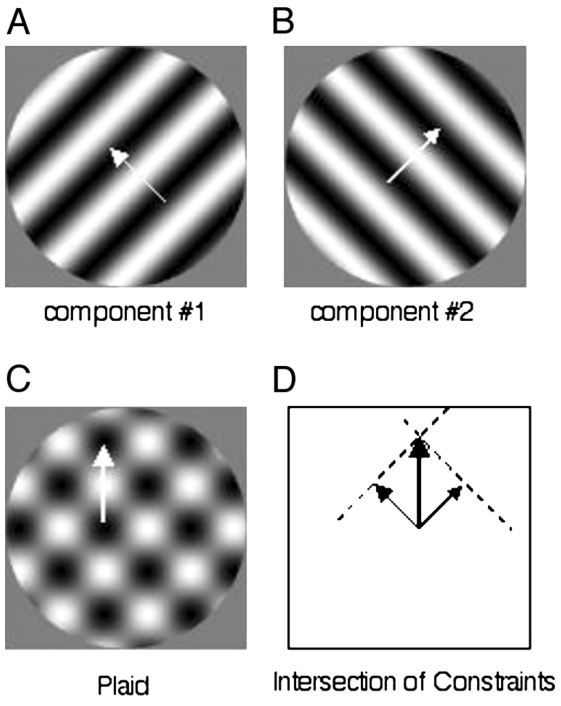

Figure 6.2: Calculation of the IOC direction (D) for a plaid (C) from its component directions (A & B). From Tinsley et al. (2003)

and directions of motion. If the component gratings are sufficiently similar in terms of spatial frequency, contrast and motion direction , the stimulus will be perceived as a coherently drifting pattern. The perceived direction of motion for this pattern will not match either of the component gratings, but can be calculated using a computation called the Intersection of Constraints (Adelson and Movshon, 1982). This model suggests a two-stage process for the computation of visual motion: the extraction of a set of 1-dimensional local motion estimates (corresponding to the motion of the component gratings), and a subsequent stage where these 1D vectors are combined via some non-linear process. Evidence for an initial stage of component motion extraction comes from the finding that masking of plaid patterns with 1D noise is most effective when it matches the orientation of the components of a plaid (Adelson and Movshon, 1982), and the finding that perceived plaid direction is based on the perceived direction of the underlying components (Derrington and Suero, 1991), both of which indicate that computing the motion of a plaid begins with estimating the motion of the components.

Evidence for area MT as the brain area in which 2D motion is computed stem from a number of neurophysiological results. Many direction-selective neurons in layer 4B of V1 project directly to MT (Maunsell and van Essen, 1983), and nearly all cells in MT are themselves selective for direction of motion (Maunsell and van Essen, 1983; Zeki, 1980). When stimulated with drifting plaid stimuli, the activity of a large number of cells in V1 could be predicted from the response of that cell to the underlying components of the plaid, resulting in a bi-lobed tuning curve with a peak (Movshon et al., 1985) (Figure 6.3 A & B). Cells with this kind of response property have been called 'component-selective' cells, due to the fact they were selective only for the motion of the component gratings, consistent with the first stage of the IOC model (Adelson and Movshon, 1982). When the tuning curves for cells in MT in response to plaids were measured, ~40% showed the same 'component-selective' responses. However, ~25% instead showed a single-lobed tuning curve in response to the overall plaid motion, consistent with the second-stage of the IOC model (Figure 6.3 C & D). The remainder of MT neurons showed 'unclassifiable' tuning curves in response to the plaid patterns. Although an IOC computation is capable of extracting pattern motion from a set of component motions, it is unclear how an MT cell could perform this computation. Simoncelli and Heeger (1998) developed a model of MT selectivity that was a neural implementation of the IOC model. Because the model involves summing the outputs of local motion filters that lie on a plane in Fourier space, it is sometimes referred to as the F-Plane model. This model can account for a number of properties of pattern-selective cells in MT, including some seemingly surprising findings such as bimodal tuning for bars moving at slow speeds (Okamoto et al., 1999). Evidence consistent with

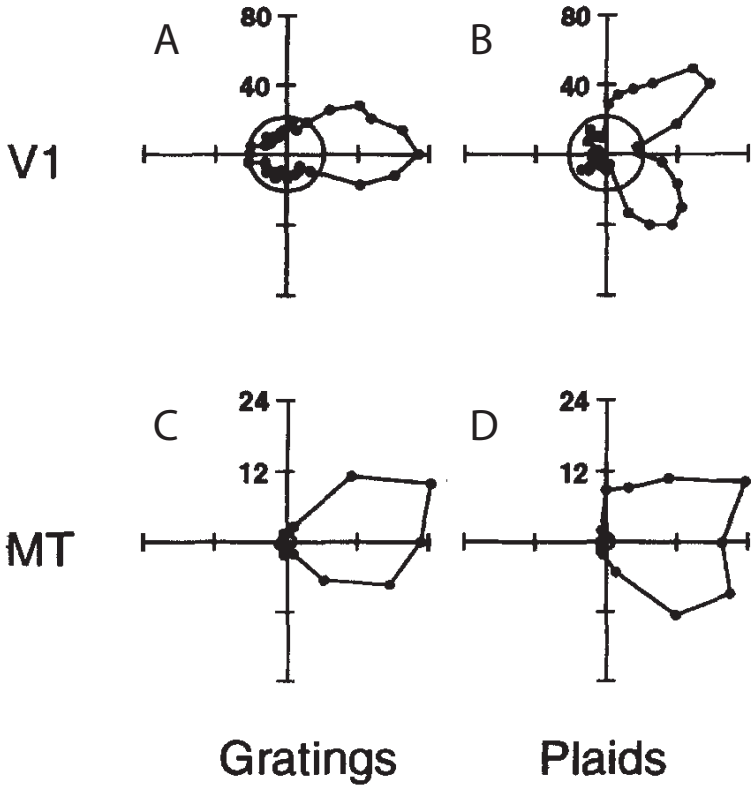

Figure 6.3: Component-motion selectivity and pattern-motion selectivity. (A, B, C, D) Direction tuning curves of a component selective V1 neuron and a pattern motion selective MT neuron. (Simoncelli and Heeger (1998) re-plotted from Movshon et al. (1985)). Stimuli are drifting gratings, and plaid patterns composed of two gratings. Response is plotted radially and the direction of stimulus motion is indicated by the angular coordinate. Circles indicate the spontaneous firing rates. The direction tuning for plaids is bimodal in the V1 neuron, indicating that these neurons respond separately to the motions of the two component gratings. The direction tuning curves for plaids are unimodal in the pattern-selective MT neuron, indicating that neurons of this type respond to the combined motion of the entire plaid pattern, not to the motions of the component gratings.

an IOC or F-Plane model indicates that motion of a 2D pattern such as a plaid is not available to the visual system prior to MT, and is computed exclusively in this area.

An alternative account of how processing of object motion may occur in visual cortex comes from experiments involving the motion of features of the pattern (corners, dots etc.). A particular stimulus used to study this is an oriented bar, with a direction of drift non-orthogonal to its orientation. The component motion for such a stimulus is orthogonal to the line's orientation, whereas the ends of the line move in the veridical pattern motion. The ends of the line, or terminators, represent features, which can be used to accurately measure the motion of the pattern. When tested with stimuli of this type, neurons in MT showed an initial bias towards motion perpendicular to the bar's orientation, with later responses (~80ms later) instead being biased towards the true pattern motion as signalled by the features (Pack and Born, 2001). One key feature of this result was that this held for the vast majority of MT cells, with no split into 'component-' or 'pattern-selective'. A similar delay has also been shown for the development of pattern-selectivity for plaids in pattern cells in MT (Smith et al., 2005), so such a result is potentially also consistent with a delayed computation of the IOC rather than the use of features in pattern-motion computation. Results from experiments using stimuli with explicit features such as drifting bars do however point to features of moving patterns as a potential alternative to two-stage computations such as the IOC.

One potential mechanism for solving the aperture problem is end-stopping in direction-selective cells (Hubel and Wiesel, 1965), where cells reduce their response to a contour that extends beyond the cells RF. Cells such as these should be capable of signalling the true 2D motion of a feature such as a line terminator, and cells such as these have been demonstrated in monkey V1 (Pack et al., 2003), with the end-stopping effect also taking ~80ms to develop, indicating that it may play a role in the pattern-selectivity demonstrated in MT (Pack and Born, 2001). Although it has not been directly demonstrated that end-stopped cells in V1 project to MT, given that ~90% of the input to MT from V1 originates from layer 4B, and that cells in this layer of V1 are known to be highly direction-selective and end-stopped, it would seem likely that these inputs possess these qualities. If the 2D motion of pattern-features can be signalled directly by end-stopped cells in V1, this means that the aperture problem can be solved in V1, meaning that the role of cells in MT is potentially to simply pool the outputs from V1, rather than compute pattern motion.

Although end-stopping may explain pattern-selectivity in MT for tilted bar stimuli, it is not clear whether it can explain selectivity for patterns such as plaids. The overlap re-

gions of the plaids constitute candidate features that could be tracked by end-stopped neurons in V1. Although initial reports reported no pattern-selectivity in V1 (Movshon and Newsome, 1996; Movshon et al., 1985), a number of V1 cells with 'pattern-selective' like responses to plaids have been found in monkey V1 (Guo et al., 2004; Tinsley et al., 2003), providing further evidence for the possibility that the processing of pattern motion may not solely be performed in higher visual areas. Similarly to findings in MT, Tinsley et al. (2003) showed that the full range of pattern-responses, from purely component-selective to pattern-selective could be found in V1, depending on the size and shape of the cells receptive fields. Whilst only a small number of these cells could be classified as purely 'pattern-selective', they indicated that a pattern-motion signal was available in V1. In addition, anaesthesia abolished pattern-responses in V1 (Guo et al., 2004), an effect that has also been observed for some pattern-responses in MT (Pack et al., 2001), although this result has been disputed (Movshon et al., 2003).

To what extent 'true' pattern-selectivity exists in V1, and what contribution it makes to pattern-selectivity in MT, has not been directly demonstrated, and it has been argued that what evidence exists is unlikely to fully explain the 'pattern-motion' responses seen in MT under circumstances when no 'pattern-motion' responses can be seen in V1 (Movshon et al., 1985). It may be the case that a two-stage process involving linear filtering in V1 followed by a nonlinear computation in MT coexists with a feature-tracking process in V1, with the two processes trading dominance under different conditions (Weiss et al., 2002). Whilst end-stopped cells dominate layer 4B of V1, layer 6 also projects to MT and has largely end-free direction-selective cells, indicating a second possible route for component-motion information to each MT, one that competes with feature-motion via layer 4B.

Evidence for pattern-selectivity in the human visual cortex has been shown using fMRI. Huk and Heeger (2002) examined fMRI adaptation effects in MT+ for plaid patterns where the pattern motion was kept constant as opposed to where the pattern motion varied. In both cases, the exact same range of underlying component motions were used, so any adaptation effects observed would be due to the pattern motion itself. They found reduced BOLD signal in MT+ during the constant pattern motion blocks, that was not apparent during the blocks with varying pattern motion. They did not find significant adaptation in V1, and found varying levels of adaptation in intermediate visual areas. When the perception of coherent pattern motion was removed by inducing transparency (by changing the spatial frequency of the component gratings), the adaptation effects disappeared, indicating that is in fact selectivity for coherent pattern motion being measured in MT+. In intermediate visual areas between V1 and MT

(such as V2, V3, V4 and V3A) increasing amounts of pattern-motion adaptation was found, indicating a possible role for these areas in pattern-motion processing.

Pattern selectivity has also been demonstrated in MT+ by contrasting responses for plaid patterns to those for transparent grating surfaces, with findings of larger responses for the component-motion percept compared to the pattern motion percept, possibly due to the fact that the multiple component motion percepts activated multiple neural populations whereas the single coherent plaid percept activated only one (Castelo-Branco et al., 2002; Villeneuve et al., 2005, 2012). This pattern of responses was not seen in V1 (Villeneuve et al., 2005, 2012), indicating that this area does not have a role in pattern motion processing, and as with Huk and Heeger (2002) intermediate visual areas also showed pattern-motion specific effects, indicating a possible distributed network for pattern motion processing in human visual network.

In the experiments described in this chapter, classifiers were trained based on the patterns of fMRI responses elicited by drifting gratings and plaids formed from gratings. To test for pattern-motion selectivity, i.e. areas that responded to the motion of a pattern regardless of its underlying motion components, classifiers trained on one class of stimuli were tested with stimuli of the other class. The logic of the experiment was as follows: if an area showed 'component-selective' responses, the classifiers would be unable to generalise as the component motions of grating and plaid stimuli with the same overall direction of drift would not match. If however an area showed 'pattern-selective' responses, then the classifier may be expected to generalise between stimuli with the same direction of drift, even if the underlying components were different. Given the response properties shown in macaque V1 and MT, we would expect to see no generalisation in V1, but significant generalisation across stimulus types in MT+.

6.1 Grating to Plaid Classification

6.1.1 Methods

Participants, Procedure and Stimuli

Three experienced subjects took part in the study, in two separate scanning sessions with different stimulus parameters.

Visual Stimuli Stimuli consisted of drifting, oriented square wave gratings (Figure 6.4, Panel A) (spatial frequency of 0.5 cycles/degree), or plaids formed from square wave gratings. The gratings were presented in an annulus. In the grating condition, gratings

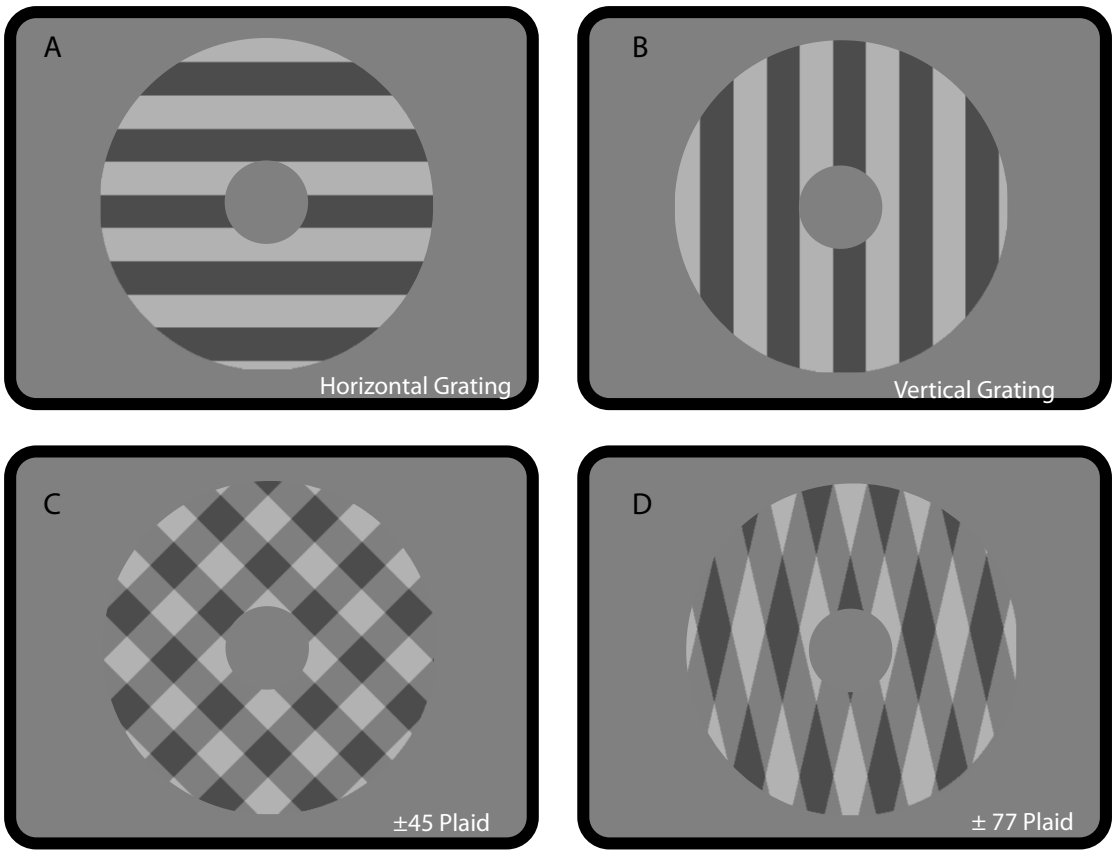

Figure 6.4: Visual stimuli used in the grating/plaid classification experiment. Square wave gratings oriented at 0 or 90 were used in the gratings scans (Figure 6.4 A). For the plaid scans, one session used orthogonally oriented gratings (Figure 6.4 B), while the second session used plaids with a component separation of 144° degrees (Figure 6.4 C and D).

had a contrast of 0.4, and in the plaid condition a contrast of 0.2, to match the maximum contrast in the stimuli across conditions. In both sessions, gratings had orientations of 0 or 90, with either positive or negative phase shift, leading to 4 possible directions of motion. In the first scanning session, plaids were formed from gratings oriented +45° and -45° to the desired direction of motion, forming the plaid pattern shown in Figure 6.4, Panel B. Plaids in the second session were formed from gratings oriented +77° and -77° to the desired direction of motion, forming the plaids seen in Figure 6.4, Panels C and D. These plaids were used to test whether generalisation was dependent on the angular separation between the components.

As a follow-up experiment, the $\pm 77°$ session was repeated for two subjects, but using sinusoidal gratings to form the stimuli used in the grating and plaid stimuli, to avoid issues of stimulus transparency in square wave plaids with a large angular separation.

Scanning Session Grating and plaid scans were interleaved within each scanning ses-

sion. Stimuli were presented in a block design, with 16s per direction, each direction being shown twice per scan.

Localiser Scan Two motion localiser scans were performed per session, to identify motion sensitive areas that responded to the visual stimulus. The motion localiser stimulus consisted of a cloud of dots that alternated between static and motion (Huk et al., 2002). In addition, two motion retinotopy scans were performed in each session(Huk et al., 2002).

Attention Control Task To control for changes in the attentional state of observers, they were asked to perform a demanding contrast discrimination task at fixation, as described in Chapter 2.

Data Analysis

For each scanning session, MVPA and leave-one-run-out cross validation was performed as described in Chapter 4. In addition, to test generalisation the classifiers trained on all but one of the grating runs were tested on a single plaid run, and vice versa. Leave-one-run-out cross-validation was used to ensure that the same data were not used for testing and training, and to equalize the amount of training data in the classification and generalisation conditions. Classification analysis was performed in two complementary ways: 1) on the basis of ROIs (selecting 500 voxels from each on the basis of a motion localiser), and 2) using a spherical searchlight method.

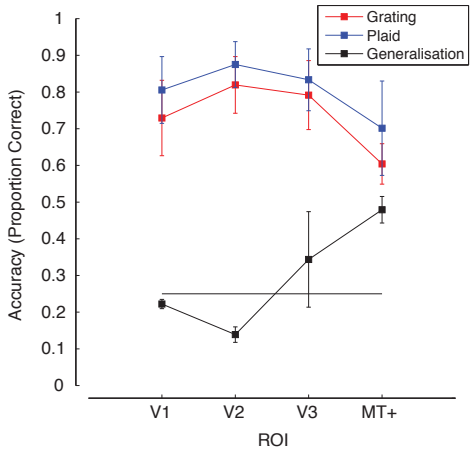

Figure 6.5: Leave one run out cross validation results for the ±45° plaid sessions across 3 subjects. Error bars indicate ±1 SEM (n=3).

6.1.2 Results

Classification for plaids and gratings (Figure 6.5 red and blue Lines) for ± 45° sessions was above chance in all areas tested, as tested by a within subject permutation test. The generalisation (Figure 6.5 Black Line) was tested by taking the classifiers trained on the plaid and grating stimuli and testing them on the stimuli of the opposite class. V1 and V2 showed no above-chance generalisation, whilst MT+ showed above-chance performance as measured by within subject permutation tests. V3 showed above-chance performance for some subjects. Generalisation results for individual subjects are shown in Table 6.1. Results from a searchlight classification analysis (Figure 6.6) reiterate the result from the ROI classification, showing elevated accuracies across visual cortex, whilst generalisation only shows elevated accuracies in the vicinity of MT+ on the lateral surface.

	V1	V2	V3	MT
ab	0.198 (0.798)	0.135 (0.961)	**0.375 (0.017)**	**0.417 (0.007)**
ms	0.188 (0.855)	0.104 (1.000)	0.104 (0.999)	**0.479 (0.001)**
rs	0.229 (0.568)	0.177 (0.874)	**0.552 (0.001)**	**0.542 (0.001)**

Table 6.1: Generalisation results for ±45° plaid session showing mean accuracies for individual subjects and ROIs. Figures in brackets show the statistical significance of each accuracy score based on a permutation test.

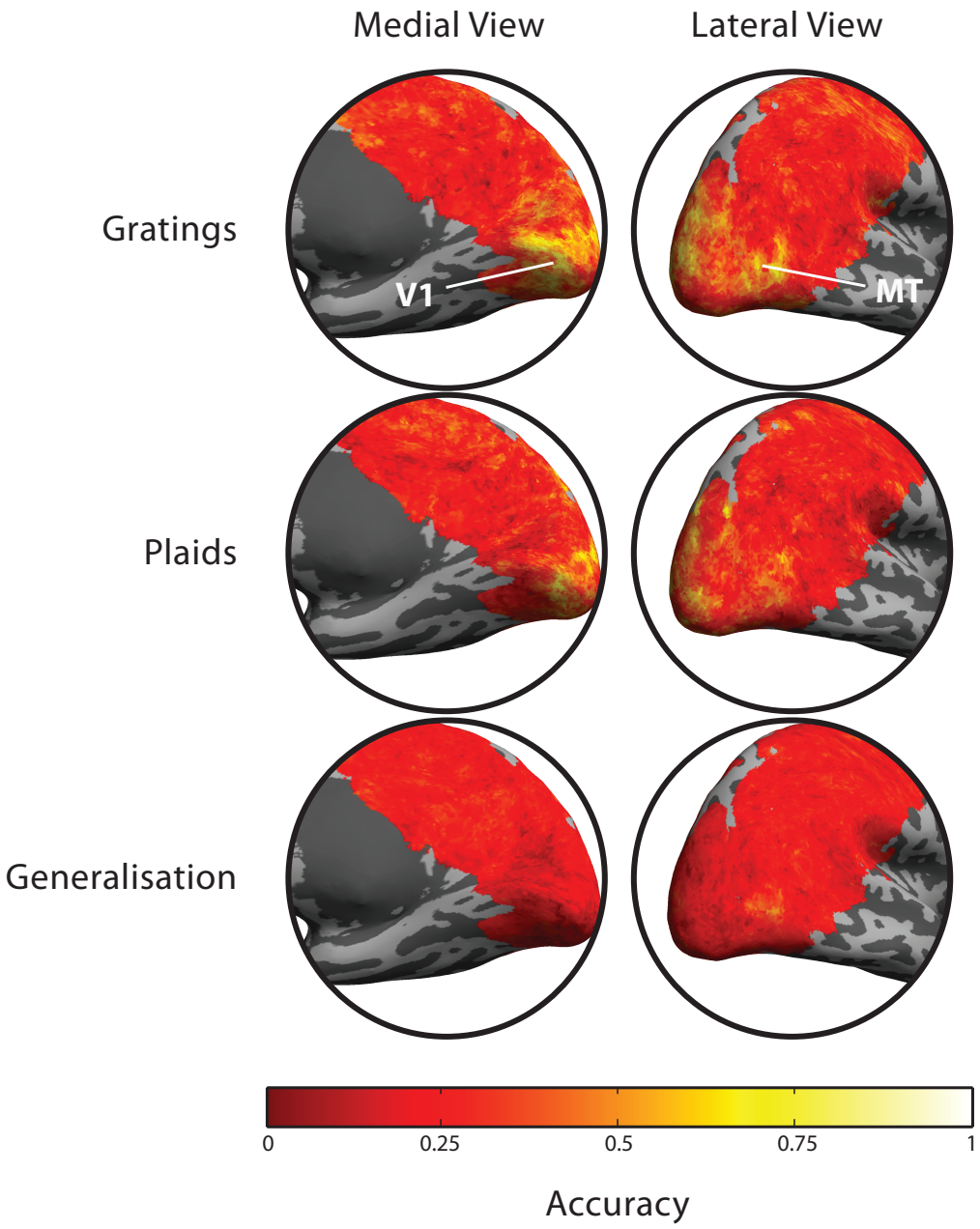

Figure 6.6: Accuracies from the searchlight classification analysis for a single subject displayed on the inflated cortical surface. Grating and plaid classification (Top 2 Rows) show elevated classification accuracies across visual cortex, including MT+. Generalisation only shows elevated accuracies in MT+. Chance performance = 0.25.

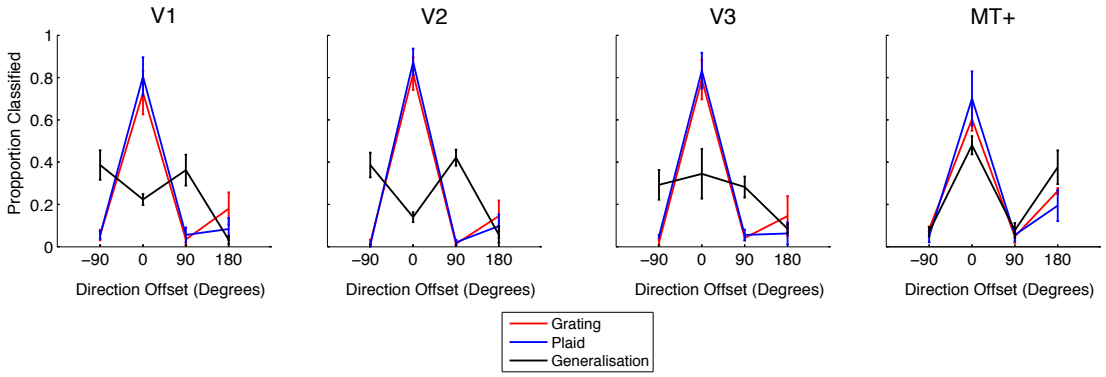

Figure 6.7: MVPA tuning curves for the $\pm 45^{\circ}$ plaid session, indicating how often each direction of motion was classified as one of the 4 possible directions. Error bars ± 1 SEM (n=3)

In a more detailed analysis, the MVPA tuning curves for the grating and plaid classification were examined (Figure 6.7 red and blue Lines). A a similar pattern of results to those found for random dot stimuli (Chapter 4, Kamitani and Tong (2006)), with a peak at the correct direction, very few misclassifications in the orthogonal directions, and some misclassifications in the opposite directions. The tuning curves for the generalisation condition varied depending on whether successful generalisation was found for that area. In V1 and V2, the curve peaked at the orthogonal directions, with the curves flattening out in higher areas, before the curve in MT+ returned to the pattern seen in the grating and plaid classification.

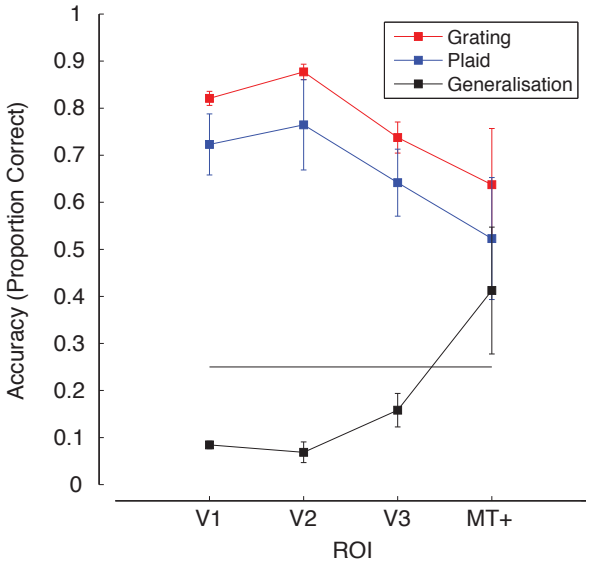

Figure 6.8: Leave one run out cross validation results for the $\pm 77^\circ$ plaid sessions across 3 subjects. Error bars indicate ± 1 SEM (n=3).

For the $\pm 77^\circ$ plaid sessions, classification for the plaids and gratings was above chance in all areas examined (Figure 6.8, red and blue Lines). The tuning curves also matched those from the $\pm 45^\circ$ session (Figure 6.9, red and blue lines). The generalisation did not show consistent above chance performance in any visual area, including MT+, although accuracy was higher in this area. Results for individual subjects (Table 6.2) showed a substantial variability in classification accuracies, from well above chance to near chance levels in MT+.

	V1	V2	V3	MT
rs	0.100 (0.988)	0.087 (0.993)	0.200 (0.750)	**0.662 (0.001)**
ms	0.078 (0.996)	0.094 (0.986)	0.188 (0.807)	**0.375 (0.016)**
ab	0.075 (0.999)	0.025 (0.999)	0.087 (0.996)	0.200 (0.795)

Table 6.2: Generalisation results for $\pm 77^\circ$ plaid session showing mean accuracies for individual subjects and ROIs. Figures in brackets show the significance of each value based on a permutation test.

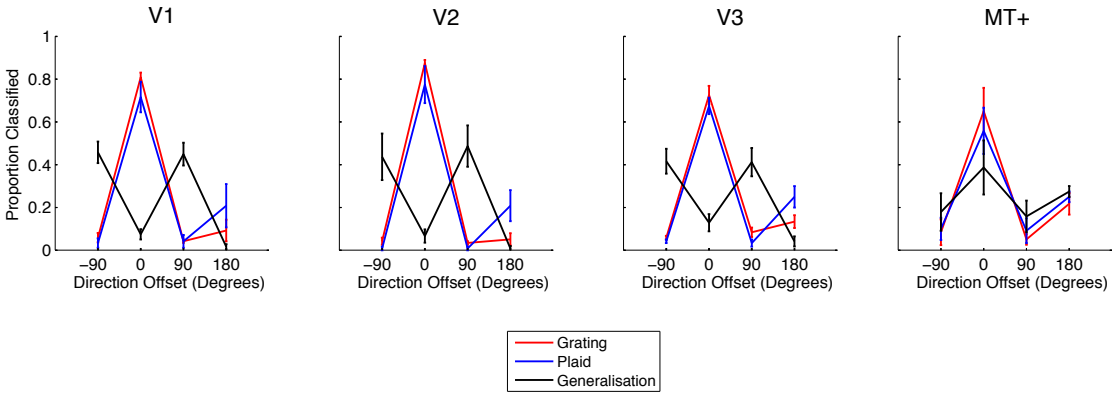

Figure 6.9: MVPA tuning curves for the ±77° plaid session, indicating how often each direction of motion was classified as one of the 4 possible directions. Error bars indicate ±1 SEM (n=3)

The MVPA tuning curves for the ±77° plaid session (Figure 6.9) are similar to those from the ±45° session, with classification for both gratings and plaids showing a peak at the correct direction followed by a smaller peak for the opposite direction of motion (Figure 6.9, red and blue lines). For V1-V3, the peak at the orthogonal directions of motion for the generalisation curves (Figure 6.9, black lines) are even more pronounced than those seen in the ±45° session, indicating that when trained with stimuli from the other class, stimuli would more frequently be classified as orthogonal directions. For MT+, the tuning curve more closely resembles the curves for grating and plaid classification than the generalisation curves for V1-V3.

One possible reason for the failure to find significant generalisation in MT+ for the ±77° plaids is the reports of transparency using the ±77° square-wave plaids. As transparency would be expected to interfere with the generalisation, the ±77° experiment was repeated with two subjects, but using sinusoidal gratings for the stimuli. Generalisation results were again mixed, with classification being higher in MT+. The results for the individual subjects (Table 6.3) showed that one subject showed significantly above chance performance, whilst the other subject failed to meet significance, although performance was greater than chance. The tuning curves for the ±77° sinusoidal plaids again showed the switch from MVPA tuning curves for generalisation (Figure 6.11, Black Lines) with peaks at the orthogonal directions in V1-V3, to a curve with a peak at the correct direction in MT+.

	V1	V2	V3	MT
ab	0.050 (1.000)	0.025 (1.000)	0.062 (1.000)	**0.500 (0.001)**
ms	0.047 (0.999)	0.016 (1.000)	0.156 (0.940)	0.328 (0.071)

Table 6.3: Generalisation results for ±77° sinusoidal plaid session showing mean accuracies for individual subjects and ROIs. Figures in brackets show the significance of each value based on a permutation test.

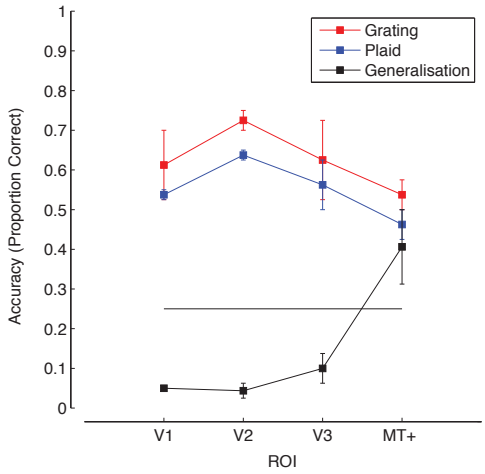

Figure 6.10: Leave one run out cross validation results for the ±77° sinusoidal plaid sessions. Error bars indicate ±1 SEM (n=2).

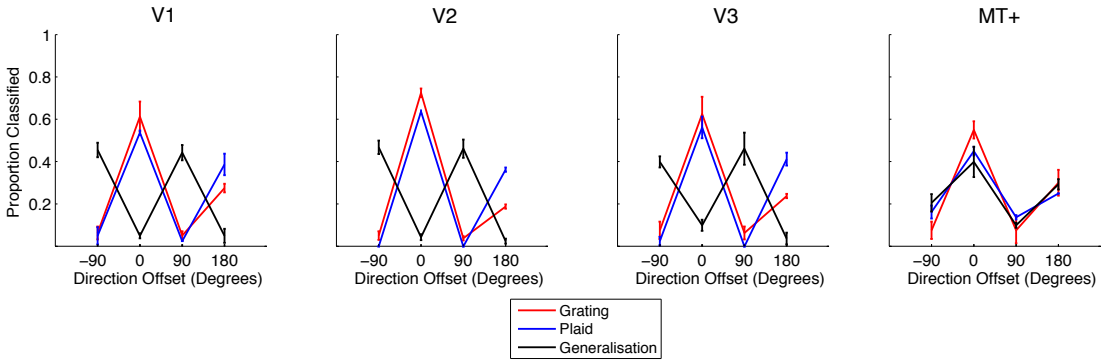

Figure 6.11: MVPA tuning curves for the ±77° sinusoidal plaid session, indicating how often each direction of motion was classified as one of the 4 possible directions. Error bars indicate ±1 SEM (n=2)

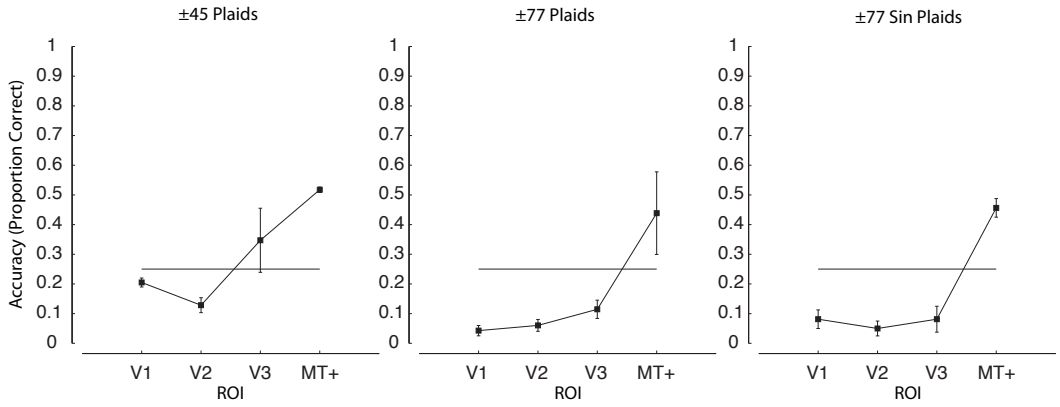

Figure 6.12: Classification results for generalisation between gratings and plaids when all the data from one stimulus type was used to train the classifier, which was then tested with all the runs of the alternate type. Error bars indicate ± 1 SEM (n=3,3 & 2).

As well as using leave-one-run-out cross validation, we also tested generalisation by training a classifier on all the grating runs and then classifying the plaid runs, and vice versa (Figure 6.12). This allowed a greater amount of data to be used to build the classifiers whilst still allowing the independence of the training and test data, given that they were drawn from completely separate runs. Results using all runs mirrored those for 'leave-one-run-out' cross validation, with higher generalisation performance in MT+ for $\pm 45^\circ$ plaids, with more variable performance for $\pm 77^\circ$ plaids. Using all runs for test and training shows good performance in the $\pm 77^\circ$ sinusoidal grating plaids condition. Results for the 2 individual subjects also showed significance in MT+ as measured by a permutation test (Table 6.4).

+/- 45	V1	V2	V3	MT
ab	0.229 (0.597)	0.094 (0.994)	**0.417 (0.005)**	**0.500 (0.001)**
ms	0.135 (0.986)	0.115 (0.995)	0.135 (0.990)	**0.521 (0.001)**
rs	0.208 (0.676)	0.177 (0.854)	**0.490 (0.002)**	**0.531 (0.001)**

+/- 77	V1	V2	V3	MT
ab	0.038 (1.000)	0.025 (1.000)	0.075 (0.998)	0.225 (0.605)
ms	0.075 (0.995)	0.062 (0.998)	0.175 (0.841)	**0.700 (0.001)**
rs	0.016 (1.000)	0.094 (0.988)	0.094 (0.991)	**0.391 (0.015)**

+/- 77 SG	V1	V2	V3	MT
ab	0.050 (1.000)	0.025 (1.000)	0.038 (1.000)	**0.487 (0.001)**
ms	0.047 (1.000)	0.047 (0.999)	0.125 (0.979)	**0.375 (0.011)**

Table 6.4: Generalisation results for each plaid session using all runs to train/test showing mean accuracies for individual subjects and ROIs. Figures in brackets show the significance of each value based on a permutation test.

6.1.3 Discussion

We investigated 'pattern selectivity' in MT+ by training and testing classifiers using the voxel patterns resulting from drifting plaids and gratings. In the initial experiment the plaids were formed from $\pm 45^{\circ}$ oriented drifting gratings, whilst the gratings were horizontally or vertically oriented. In both cases, the overall drift directions were matched. Above chance classification was seen for both classes of stimuli in early visual areas (V1-V3) and MT+, matching the results seen using random dot stimuli (Kamitani and Tong (2006) , Chapter 4). Examining the 'tuning functions' for these classifications revealed a similar pattern of responses as seen using dot stimuli: a peak at the correct direction, with very few misclassifications in the orthogonal directions, and occasional misclassifications in the opposite direction.

When the classifiers were tested on the the opposite class of stimuli, early visual areas did not show above-chance performance, whereas the classifiers trained on voxels from MT+ were above chance, showing that the patterns of activity elicited by plaids and gratings with the same overall drift direction were similar enough a classifier trained on one class of stimuli could accurately classifier the other. This result is in line with the known tuning properties of pattern-selective cells in macaque MT, and with previous fMRI results indicating human MT+ as an area selective for pattern-motion of plaids (Huk and Heeger, 2002; Villeneuve et al., 2012).

Note that although we found no evidence of grating to plaid generalisation in V1, this does not preclude some kind of pattern selectivity in V1. Successful generalisation relies on the patterns elicited by a drifting grating moving in a given direction being similar enough to the patterns of a drifting plaid moving in the same direction. It may be the case that the pattern selective cells in V1 selective for a given plaid direction differ from those selective for a grating moving in the same direction, for example in terms of spatial frequency tuning. Nevertheless, we were able to demonstrate generalisation from gratings to plaids in MT+ in situations where none was seen in V1, indicating pattern selectivity in MT+.

Although selectivity for pattern-motion is one explanation for the generalisation seen in MT+, an alternative is that the tuning of direction-selective cells in MT+ is simply broader than those in early visual areas, a known property of cells in macaque MT (Albright et al., 1984). If direction selective cells have broad enough tuning, the tuning curves for the components may overlap, yielding a single broad tuning curve centred on the pattern-motion of the plaid (Tinsley et al., 2003). It could be that in $\pm 45^{\circ}$ plaids, the underlying component motions are close enough to the overall pattern motion that

they activate component-selective cells tuned to that direction, allowing generalisation even without selectivity for pattern motion.

To test this possibility, we repeated the experiment with plaids formed from components oriented $\pm 77^{\circ}$ relative to the pattern motion. In plaids such as these, the underlying components should not be close enough to the predicted pattern-motion direction to activate any component-selective cells tuned to this direction, meaning that any generalisation could not be explained in this fashion. When tested with these stimuli, above chance classification was found in MT+ for 2 of the subjects, whilst the remaining subject showed only chance performance. In debriefing, this subject reported that some of the plaid stimuli appeared as transparent gratings rather than a coherently drifting pattern. Plaid stimuli of this type, square wave gratings with a large angular separation, do have the potential to appear transparent even when the components are matched in terms of spatial frequency (Stoner and Albright, 1996), and such transparency would be expected to interfere with successful generalisation, as motion transparency causes pattern-motion cells to 'revert' to component-motion responses (Huk and Heeger, 2002).

To test whether this was the case, two subjects were tested using $\pm 77^{\circ}$ plaids formed from sinusoidal, rather than square-wave, gratings. Square-wave gratings were initially used to make the stimuli more spatially broadband, but as stated above this increases the chance of the plaids being perceived as transparent rather than coherent. Plaids formed from sinusoidal gratings are more likely to be perceived as coherent, and generally require the spatial frequencies of the underlying components to be changed in order to induce transparency (Huk and Heeger, 2002; Villeneuve et al., 2012). No subjects reported any perception of transparency for the sinusoidal stimuli, and above-chance generalisation was found for both subjects, albeit only when all data was used for classification and testing, based on a permutation test. This confirms that classification can generalise between plaids and gratings with the same overall direction of drift, even when the underlying plaid components are very different from the perceived pattern direction.

Overall, the results suggest that generalisation between the classification of plaids and gratings is possible using patterns of voxel activity from MT+, but not in earlier visual areas, and that this generalisation is dependent on the plaid stimuli being perceived as coherently drifting stimuli. This indicates that the patterns produced by drifting gratings and plaid stimuli with the same overall direction of drift are similar enough that a classifier trained on one set of stimuli can successfully classify the other set, indicating that a similar population of cells are activated by similar directions of drift in the

two stimulus classes (Harrison and Tong, 2009; Hong et al., 2011). This is consistent with previous fMRI results, which showed the strongest fMRI pattern-motion in MT+, indicating a large proportion of pattern-motion sensitive cells in this area (Huk and Heeger, 2002). The ability for areas with a high proportion of pattern-selective cells to generalise between plaid and grating stimuli is consistent with the response properties of such cells in the macaque, which showed a similar tuning curve for motion in that cells preferred direction, whether the stimulus was a drifting grating or a coherently drifting plaid pattern (Albright et al., 1984). Areas that are more strongly component selective, such as V1, would be unlikely to show generalisation, as the directions of the underlying components do not match when comparing the grating and plaid stimuli. Intriguingly we also found below-chance classification in the $\pm 77^{\circ}$ condition consistently across subjects for area V1, which would also points to this area having a 'component-motion' like response, due to the similarity of the component motions for $\pm 77^{\circ}$ plaids to gratings moving in the orthogonal direction.

Although our results do not indicate pattern-selective responses in early visual areas, this does not prove that no patten-selectivity exists in these areas, simply that no signal in this area could be used by the classifiers used in this experiment. Pattern-selective type responses have been demonstrated in V1 neurons, especially in cells with end-stopped receptive fields that are known to project to MT (Pack et al., 2003), or in cells with short, wide receptive fields (Tinsley et al., 2003). The overall proportion of pattern-selective cells is believed to be much higher in MT, and this may be the reason that generalisation was seen in MT+ and not in V1, or it may be that while there exist cells in V1 capable of signalling pattern-motion, these cells do not respond the same way to drifting plaids and gratings.

Our results contrast with previous fMRI adaptation results indicating pattern-motion responses in additional visual areas such as V2 and V3. Huk and Heeger (2002) found pattern-motion adaptation in a range of visual areas, and pattern-motion selective responses were found in a similar range of areas when probed with coherent and transparent plaids (Villeneuve et al., 2012). We saw generalisation only in MT+, which could be accounted for by pattern-motion selective cells being present in lower-areas, but at insufficient proportions to allow a strong-enough pattern-motion selective signal that can be used by the classifier, or having response properties that do not allow generalisation.

These results seem to confirm that there exist cells in human MT+ that are selective for a given direction of motion, regardless of the underlying orientations that give rise to such a signal. This is consistent with previous fMRI results, which found similar re-

sponses to motion stimuli even when the overall spatial frequencies of the stimuli were changed. Huk and Heeger (2002) found plaid adaptation even when the plaids used had different underlying component grating orientations, and between plaids with different spatial frequency components, and Villeneuve et al. (2012) found similar responses in MT and MST to high and low spatial frequency plaids. One potential test for this would be to test for generalisation between stimuli with very different orientation and spatial frequency make-up, for example testing for generalisation between high and low frequency plaid patterns, or generalisation between $\pm 45^{\circ}$ and $\pm 77^{\circ}$ plaids.

We hypothesised that the failure to see significant generalisation for the $\pm 77^{\circ}$ square-wave plaids was due to the stimuli being perceived as transparent. This would be consistent with previous studies which found pattern-selective responses only when coherent plaids were used (Huk and Heeger, 2002; Villeneuve et al., 2012). To be sure that generalisation requires coherent plaid stimuli, generalisation between gratings and coherent and transparent plaids would need to be directly compared, using plaids where transparency could be controlled directly. Transparency in square-wave plaids can be controlled by changing the contrast of the features caused by the intersection of the gratings (Stoner and Albright, 1996), but this has the potential to introduce additional components to the plaid (Movshon et al., 2003) meaning that using this method to control transparency may not be useful in comparing coherent and transparent plaids. An additional method of inducing transparency is to use components with non-identical contrasts or spatial frequencies (Adelson and Movshon, 1982), the latter of which was used to demonstrate a lack of pattern selective responses in previous fMRI studies (Huk and Heeger, 2002; Villeneuve et al., 2012). Although this method can induce transparency in both square-wave and sinusoidal plaids, it is unclear how altering the spatial frequency of the stimuli will affect generalisation. It may be preferable to use stimuli where transparency and coherence can be controlled without altering the spatial frequencies of the stimuli.

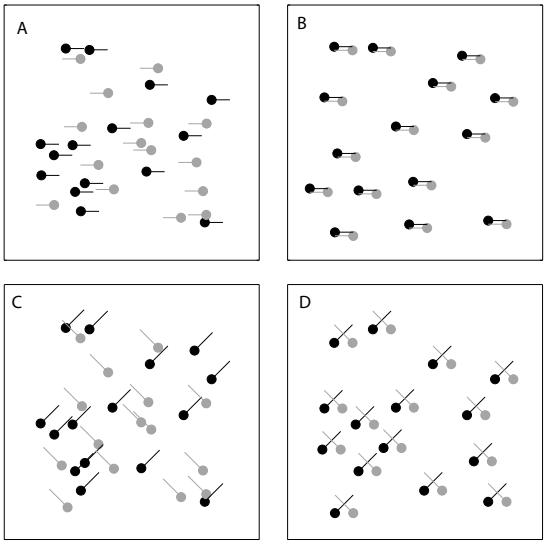

Figure 6.13: Schematic of different dot distributions leading to transparent and non-transparent motions. Two populations of randomly placed dots will appear as two transparent surfaces sliding across each other, whether the two directions are opposite (A) or orthogonal (C). If one dot from each direction is paired with a dot from the other direction, the pattern will either appear as directionless flicker (B) or a coherently drifting pattern with motion in the Vector Average direction (D).

6.2 Dot to Paired Dot Classification

As well as gratings and plaid stimuli, a particular class of random dot stimuli can also be used to investigate the computation of pattern motion in the visual system. If two drifting dot clouds with sufficiently differing directions or speeds are overlaid, they will be perceived as two transparent surfaces 'sliding' across one another, as opposed to a coherent whole. This transparency can be removed, however, by changing the way in which the dots are arranged relatively to each other. If each dot in a coherently moving surface is paired with a dot moving in the opposite direction, such that they move across each other over a short distance, the perception of transparency is removed (Qian et al., 1994) (Figure 6.13). These 'paired' dot patterns are perceived as directionless flicker, with no perception of the opposing motions, whereas unpaired dots with the same distribution of directions are perceived as two transparent surfaces moving in opposite directions. This motion cancellation only occurs if the dots are sufficiently close and the dot pairs travel only a short distance, indicating that the mechanism by which the two dot directions interact is a local one with a small-scale.

Some cells in macaque MT have been shown to modulate their response to paired

versus unpaired dot patterns. Cells in MT fire strongly when dots drift in the cells preferred direction, and suppress their firing when dots drift in their anti-preferred direction. If they are presented with two populations of dots drifting in the cells preferred and anti-preferred directions respectively, the cell will fire, but not as strongly as for dots drifting solely in its preferred direction, due to the suppressing actions of the anti-preferred direction. If the dots were paired so that only the flicker was seen, the firing rates of ~40% of cell tested were additionally suppressed (Qian and Andersen, 1994). Although this modulation suggests that cells in MT are the locus for the perception of transparency, the maximum separation of dots in a pair before the perception switches from flicker to transparency is ~0.4°, which is far smaller than the receptive field of cells in MT. This means that either the interactions between paired dots occur within MT 'subunits' with smaller RFs, or possibly within V1 direction-selective cells themselves. However, little suppression of activity for paired dot patterns compared to unpaired patterns was found in V1 (Qian and Andersen, 1994, 1995), suggesting that it is in fact MT where transparency is calculated.

FMRI in humans has also implicated MT+ as the region where these interactions are computed. Heeger et al. (1999) measured BOLD responses to paired and unpaired stimuli, and found reduced responses in MT+ for paired when compared to unpaired, with little to no reduction in signal in V1. This finding was also found in a number of other studies (Garcia and Grossman, 2009; Muckli et al., 2002). Garcia and Grossman (2009) also found little modulation in V1, but found modulation in a number of areas besides MT+ for paired versus unpaired, implying a much wider range of areas involved in transparency perception. They also compared responses for unidirectional and transparent dot patterns, and found reduced activity in MT+ only for the transparent patterns. They suggested that the reduction in activity as a result of motion-cancellation due to dot pairing was a general feature of the visual system, whilst the reduction in activity caused by opposing, transparent motions was a feature only of MT cells.

If paired dots with non-opposing directions are used, rather than a motion-cancelled 'flickering' pattern, the dots can appear as a coherently drifting pattern in the vector average of the two underlying dot motions (Curran and Braddick, 2000). If the same neural mechanisms responsible for the perception of transparent and non-transparent motion in opposing direction stimuli also govern the perception of coherent motion for non-oppositional paired dot stimuli, then we would expect that cells in MT+ should signal the coherent motion direction, whilst cells in V1 should be unable to do this. Furthermore, generalisation should only be possible when the dots are paired and perceived as a coherent pattern, if the dots are unpaired and perceived as transparent

surfaces, performance in MT+ should match that in visual areas with component selectivity.

To test this, the grating/plaid experiment was repeated, but this time with unidirectional dot patterns and paired dot patterns with Vector Average (VA) directions that matched the unidirectional dot patterns, but whose components did not match. The same reasoning applies here as for the grating/plaid experiment: the signals arising in MT+ from unidirectional and paired-dot stimuli with the same perceived direction should be similar enough to allow a classifier to generalise between the two stimuli types. We also tested generalisation between unidirectional patterns and patterns which had the same underlying dot motions as the paired dot patterns, but where the dots were arranged randomly, leading to transparency in these patterns. As pattern coherence is presumably required for successful generalisation, generalisation should not be possible between unidirectional dots and the unpaired dot patterns.

6.2.1 Methods

Paired dot patterns consisted of a cloud of dots in which half of the dots had the same direction of drift, with a limited lifetime defined by a limited trajectory. The dot stimuli were similar to those used in the experiments in Chapter 4, but with parameters selected to elicit the perception of paired motion (white dots on black ground, dot speed = 2.5°, dot density = 6 dots/deg^2, dot size = 0.1°). Each of these dots was paired with a dot moving in the orthogonal direction, arranged so that their trajectories crossed at their midpoint (Example given in Figure 6.13). Where dots overlapped, they occluded one another. Once each dot had reached the end of its trajectory, the dot pair was redrawn in a new location such that their trajectories would again cross. Dots had a lifetime of 76 ms (0.19 degree dot trajectory). Paired dot patterns were perceived as coherently drifting in the vector average of the two dot distribution directions. If the dots in each pair were drawn such that their trajectories did not overlap, they were perceived as two transparent dot surfaces sliding across one another.

One subject, who had also taken part in the plaid classification experiment, took part in the paired dot experiment. As with previous classification experiments, the subject performed the contrast discrimination task at fixation to maintain attentional state.

6.2.2 Results & Discussion

	V1	V2	V3	MT+
Paired	0.208 (0.801)	0.188 (0.820)	0.250 (0.432)	**0.375 (0.025)**
Unpaired	0.292 (0.191)	0.250 (0.413)	0.292 (0.232)	0.271 (0.301)

Table 6.5: Results from the permutation test for generalisation between unidirectional dots and paired and unpaired dots, using leave-out-one-out cross validation.

Figure 6.14 shows the classification and generalisation results for a single subject for unidirectional and paired dot stimuli, and for the unidirectional and unpaired dot stimuli. Above chance generalisation for the paired dots was only seen in MT+ (Table 6.5 and Table 6.6), mirroring the results seen using plaid and grating stimuli. It should be noted that successful classification of the paired dot patterns was only seen in MT+, which would also explain generalisation only being seen in this area, as successful generalisation between stimuli requires that those stimuli can be successfully classified themselves. This contrasts with the plaid results, where successful classification of plaids and gratings was seen in all areas. It may be that as random dots are an inherently noisier stimulus than gratings, in areas without pattern selectivity, the paired dot patterns were too noisy to allow classification. Classification for unidirectional dots was also not significant in V1, despite the fact that classification of dot direction has previously been demonstrated in this area (Chapter 4, (Kamitani and Tong, 2006)), which also could have interfered with potential generalisation in this area. Results for the unpaired dots did not show above-chance generalisation in any area, including in MT+ (Table 6.5 and Table 6.6), even though the paired and unpaired patterns only differed in the arrangement of the dots, with the same underlying motion signals. Although these results come from a single subject, they suggest that paired dot stimuli can be used to demonstrate pattern selectivity in the visual cortex, and that contrasting the results of paired and unpaired dot stimuli can be used to indicate true pattern selectivity.

	V1	V2	V3	MT+
Paired	0.188 (0.800)	0.146 (0.946)	0.208 (0.687)	**0.396 (0.016)**
Unpaired	0.333 (0.075)	0.271 (0.304)	0.250 (0.506)	0.250 (0.421)

Table 6.6: Generalisation results between unidirectional dots and paired and unpaired dots using all runs to train/test showing mean accuracies for individual subjects and ROIs. Figures in brackets show the significance of each value based on a permutation test.

Paired-dot stimuli are a class of stimuli where the response of cells in MT is believed to be markedly different to cells in earlier visual areas (Heeger et al., 1999; Qian and

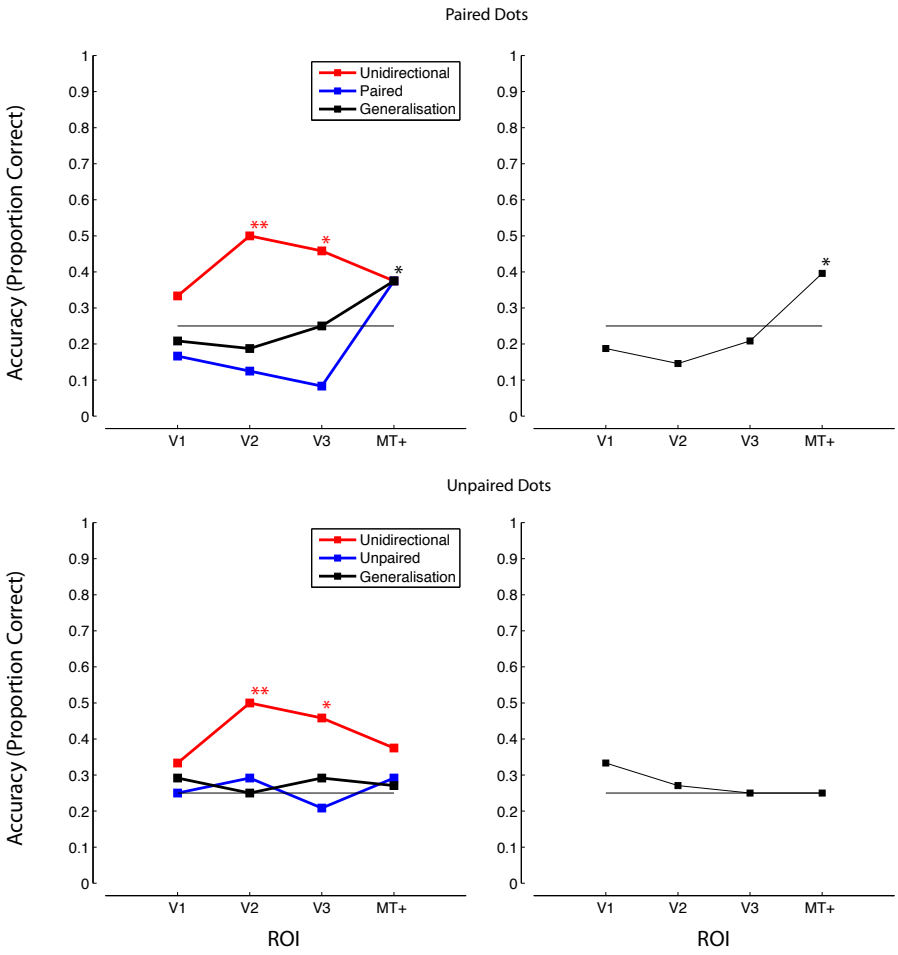

Figure 6.14: Top Row: Results for the classification of unidirectional dot patterns (red line), paired dot patterns (blue line), and generalisation between the two (black line) for a single subject using 'leave-one-run-out' cross validation (left) and training with all runs (right). Bottom Row: Results for the classification of unidirectional dot patterns (red line), unpaired dot patterns (blue line), and generalisation between the two (black line) for a single subject using 'leave-one-run-out' cross validation (left) and training with all runs (right). Asterisks indicate significance on a permutation test (* $p < .05$; ** $p < .01$; one-tailed)

Andersen, 1994, 1995), similar to the differentiation seen with stimuli such as plaids (Braddick and Qian, 2001). Cells in early visual cortex respond to the individual component motions, in this case the two populations of dots moving in orthogonal directions. Cells in MT however demonstrate an interaction between the pairs of dots with overlapping trajectories, either suppressing their activity in the case of opposite directions or responding to the vector average of the dot motions in the case of orthogonal motions.

6.3 General Discussion

We used multivariate classification techniques to study the selectivity for different types of motion in different visual areas. Early visual areas and MT+ all showed above-chance classification for drifting gratings and plaids, consistent with earlier results showing successful classification of drifting dot patterns. However, testing the generalisation of the classifiers by training on one type of stimuli and testing on the other type showed above-chance performance only in MT+. Generalisation is a test of how similar the patterns of responses elicited by one type of stimuli are to stimuli of another type. Therefore above chance performance indicates that the pattern of activity in MT+ elicited by a grating with a given direction of drift is similar to the pattern of activity elicited by a plaid with the same overall direction of drift. This is consistent with finding that a proportion of cells in macaque MT are 'pattern-motion' selective, in that they are selective for the overall motion of a stimulus rather than the motion of its components.

Another class of stimuli that can contrast pattern and component motion are paired dot stimuli, where the perceived direction of the pattern is a vector average of the two dot populations. We demonstrated (albeit in a single subject) that classification could generalise between a unidirectional dot pattern and paired dot pattern with a matching VA direction, but only in MT+. This generalisation disappeared when the dots were unpaired, indicating that this generalisation was based on the pattern direction and not the components. This provides further evidence that classification can be used to demonstrate pattern selectivity in the visual cortex.

One potential way to develop this experiment further would be to test for generalisation between a wider range of stimuli, to test how selective the responses in MT+ really are, i.e. test for generalisation between plaid patterns with different component motions, between drifting gratings and drifting dots. Tests for generalisation between motion defined by first-order and second-order cues have shown generalisation in a

range of visual areas (Hong et al., 2011), indicating a similar neural population is responsible for processing motion of these types in certain visual areas. It might be expected that pattern-motion selective cells should be able to generalise between stimuli with very different component motions.

The experiments in this chapter indicate how multivariate classification results can be useful even when the signal underlying the classification result is not completely understood. Whether the successful classification of either class of stimuli is based on the distribution of columns, a complex signal via the vasculature or a coarse scale bias, the fact that classification can generalise between two different stimuli types suggests that these stimuli are processed by a similar underlying neural population. Generalisation has been demonstrated between first and second order motion (Hong et al., 2011), which indicates a shared neural substrate for these types of motion stimuli. Harrison and Tong (2009) demonstrated generalization between orientation stimuli and the patterns of activity generated in visual cortex when subjects held these patterns in working memory, indicating that orientation selective mechanisms in early visual cortex were involved in working memory. As a note of caution however, it is worth mentioning that all of these results simply indicate that the neural substrate is similar between stimulus conditions, without necessarily providing evidence as to what neural signal the classification is based on (and what scale that signal exists at).

Chapter 7

General Discussion

Since the initial discovery of BOLD fMRI, there has been a continuing debate over what exactly can be studied with this powerful, but indirect technique. The non-invasive nature of fMRI means that the neural processes involved in perception, cognition and action can be studied in humans; and the spatial resolution offered is much higher than that offered by EEG or MEG, especially using the higher field magnets now frequently available. However its indirect nature, in which neural activity is estimated via changes in hemodynamics measurable via the BOLD signal, may limit its spatial and temporal specificity. The spatial resolution of standard fMRI is too large to directly sample even the columns arranged on the cortical surface, making the study of stimulus properties organized at these scales potentially difficult. The pilot studies using stimuli of varying motion coherence presented in Chapter 3 indicate some of the problems in interpreting changes in BOLD signal when that signal arises from voxels with a mixture of response properties contained within, in that hypothesized changes in neural activity cannot always be related to an equivalent change in BOLD signal.

One potential method for measuring sub-voxel populations of neurons is the use of adaptation techniques. Differences in BOLD response to probe stimuli, either differing from or matching the adapting stimulus, following extended exposure to an adapting stimulus can tell us about the underlying neural processing for a stimulus dimension. In Chapter 3, we found reduced BOLD amplitude for a probe stimulus after adaptation with a *similar* stimulus, and less reduction in amplitude for an orthogonal stimulus. However, comparison across these stimuli in our design was complicated by the fact they had different amplitudes even in the unadapted case, highlighting the difficulties in establishing a pre-adaptation baseline response for comparison, with adapted/unadapted comparisons being made across cortical areas. Adaptation results can be be extremely sensitive to the exact parameters of stimulus and task design used,

and therefore interpretation of fMRI adaptation results can be problematic as to where these effects arise.

A potentially more powerful method of accessing subvoxel population responses is the use of multivariate techniques, where the activity of a large group of voxels is considered simultaneously. This allows, in many cases, the pattern of activities associated with certain classes of stimuli to be successfully discriminated from patterns associated with other stimuli, even when such stimuli may not be discriminable on the basis of single voxels. In the first experiment of Chapter 4 we successfully replicated the study of Kamitani and Tong (2006) showing successful classification of the direction of a cloud of drifting dots from activity in early visual cortex and MT+. Interestingly, we saw no increase in accuracy for our experiment, at 7T with 1.5 mm isotropic voxels, when compared to the original 3T experiment, suggesting that smaller voxels may not in fact give any benefit for classification. When we re-sampled our 1.5 mm data to 3 mm, we saw no drop in accuracy, and only a small drop when we further re-sampled to 6 mm isotropic. Such results cast doubt on whether multivariate methods are truly based on signals arising from neurons or columns, and support the idea that a coarser scale signal is in fact the basis of such classification results. Potential coarse scale signals include signals from large draining veins and global biases for certain directions of motion. A bias for radial orientations in V1 has been shown to be both necessary and sufficient for orientation classification (Freeman et al., 2011), and a bias for radial directions of motion has been demonstrated in this area (Clifford et al., 2009; Raemaekers et al., 2009).

In the second part of the experiment, we tested whether a retinotopically organised coarse-scale signal could explain our classification results. We averaged voxels together after binning them based on the phase of their response to a polar angle mapping stimulus, and ran the classification experiment as before. We showed that increasing the size of the bins used (decreasing the number of bins used) led to a decrease in performance, eventually dropping near chance when the largest bin sizes were used. However, when we averaged the equivalent number voxels together at random, performance dropped off much more steeply. This indicated that averaging the voxels together in a way that preserved a signal organised retinotopically, coarser than the scale of cortical columns, was sufficient for successful classification. In the third experiment of Chapter 4 we demonstrated that the coarse scale signal was not the result of a consistent mapping of certain eye movements with particular directions of motion, which would lead to a global shift in activity in retinotopic cortex for each direction.

The advantage for retinotopic (over random) averaging in early visual cortex is con-

sistent with the known bias for radial directions of motion previously demonstrated in these areas (Clifford et al., 2009; Raemaekers et al., 2009). This contrasts with MT+, which showed no such advantage for retinotopic averaging compared with equivalent random averaging. This is again consistent with a previous study that did not demonstrate a radial bias for motion in this area. This would indicate that the classification results found in this area are not due to a coarse scale signal, but instead a signal arising from the distribution of neurons/columns, and might explain the generally lower accuracies found in this areas, despite that fact that the proportion of direction selective cells in this area is believed to be higher than that in early areas.

One argument against such an explanation is that classification is also possible with stimuli that should not give rise to such a radial bias. In the orientation domain, such radially balanced stimuli include Glass patterns, the angle of which has been shown to be discriminable with multivariate classifiers (Mannion et al., 2009). In addition, the classification of rotational motion is also possible (Kamitani and Tong, 2006; Seymour et al., 2009). In Chapter 5 we demonstrated successful classification for both of these types of motion, and showed that the classification that included an orientation component could in fact be partially explained by a coarse scale bias that was preserved by averaging voxels together on the basis of the retinotopic phase. Examining the weighting for each stimulus for each voxel or super-voxel against the polar angle represented there showed a preference for a differently oriented spiral in each visual quadrant, mirroring similar results seen in response to oriented glass patterns (Merriam et al., 2012). The classification for the sign of 'pure' rotation, where any orientation information was non-informative, showed no difference between random and retinotopic averaging. This indicates that classification of this kind of stimuli may in fact be purely driven by signals with a fine-scale neural origin, or that a coarse scale bias could not be preserved by this form of averaging. This result indicates that even the classification of stimuli that should be free of potential contamination may still be in part based on a coarse scale signal. Therefore classification results should be interpreted with caution, given that a successful classification result need not necessarily arise from information arising from fine-scale neural architecture.

Regardless of the source of the signal used in classification, a successful classification result can still be useful in establishing the type of processing done by a given cortical area. The lower classification accuracies seen in MT+ during motion classification do not indicate that MT+ is less selective, but may simply indicate that the arrangement of columns on the cortical surface may not lead to sufficient voxelwise biases to allow classification, that the amplitude of the BOLD response is less than that in other visual

areas, or that a coarse scale signal does not exist to allow high classification accuracies. In Chapter 6 we used classification to demonstrate that MT+ showed 'pattern motion' selective behaviour: classification generalised between two classes of stimuli with the same overall direction drift but very different orientation components. This behaviour was not seen in early visual cortex. This demonstrates a qualitatively different behaviour in MT+ beyond simply higher or lower classification accuracies for motion stimuli. In the second experiment of Chapter 6, we demonstrated that this behaviour may extend to other stimuli that can differentiate between 'component' and 'pattern' responses, namely 'paired dot' stimuli. Successful generalisation indicates a sufficiently similar neural representation between these stimuli to allow a classifier trained on one kind to decode the other, which indicates some degree of shared neural processing for these stimuli.

Limitations and Developments

A number of potential developments could have added to the experiments performed here. We used an MT+ localizer to identify MT+ in our subjects, which fails to distinguish between hMT and hMST (Huk et al., 2002). Whilst these areas share a preference for motion stimuli, they are expected to have slightly different properties (Becker et al., 2008), which would have been missed in our experiments. We chose to only use an MT+ localizer as we wished to maximise the number of experimental scans in each session, and having an additional MST localizer would have either reduced the number of experimental scans or required subjects to be in the scanner for longer, increasing the chances of subject motion and lapses in attention. In future however, it would be illuminating to study these areas separately, particularly to see if they differ in their responses to pattern stimuli or other complex motions.

We utilized basic block design approaches to our classification studies, with extended blocks of each stimulus to ensure distinctive patterns. Although designs such as this may ensure the highest classification accuracies, they are limited in the kind of questions that can be asked. Questions based around a subject's interpretation of a stimulus, or the results of a perceptual decision, may require an event related design. Event-related paradigms have shown successful classification, and future work could involve classification of stimuli on a trial-by-trial basis.

Another potential limitation was our inability to directly track eye-movements during the scans, due to the set-up of the 7T scanner. Previous classification studies have raised the concern that consistent mapping of eye-movements to certain stimuli may

induce global shifts of activated retinotopic cortex (Kamitani and Tong, 2006), that can be exploited by a scanner without any underlying neural selectivity. Studies that have tracked eye movements have found no link between eye-movements and classification however (Freeman et al., 2011), and classification has been demonstrated using stimuli that should not lead to consistent mapping of eye-movements to stimuli (Hong et al., 2011; Seymour et al., 2009), including in experiments within this thesis (Chapter 5). However, the lack of a direct measure of eye-movements during the experiments means that such an explanation cannot be completely ruled out, and incorporating an eye-tracker into future experiments would be an extremely useful step.

It has been suggested that successful discrimination of non-opposing motion directions in drifting dot patterns by classifiers is based on the classification of oriented motion-streaks, rather than actual motion classification (Apthorp et al., 2013). The dot stimuli used in our experiments drifted at speeds above the threshold for the creation of motion-streaks, so we cannot assume that our results reflect true motion classification. To be sure that our results are not based on orientation classification, it would be necessary to repeat our experiments using slower speed stimuli, or include an orientation control within a session to examine whether orientation classification can explain the results. An alternative is to use only rotational stimuli, where any orientation information is non-informative (Apthorp et al., 2013; Kamitani and Tong, 2006; Seymour et al., 2009). However, the research questions that can be investigated with stimuli of this type may be limited, for example it would not be possible to assess the tuning of the classification.

For searchlight classification, we utilised a volumetric searchlight, where a 'spherical' searchlight is defined in Euclidean space on the 3D anatomy. Due to the folded nature of the cortical surface, this means that areas of cortex that are non-adjacent on this surface surface may be included within the 3D searchlight. A potentially more selective methodology uses a 2D searchlight defined on the measured cortical surface , and has been shown to be more spatially specific in defining informative areas (Chen et al., 2010; Oosterhof et al., 2010). The distortions in the 7T EPI images mean that searchlights defined on the cortical surfaces calculated from T1 weighted images at 3T (Chapter 2) may not select the voxels that correspond to these areas in the EPI images. It may be possible, however, to utilise the non-linear alignment process we used to convert the retinotopically defined ROIs from the undistorted anatomy space to EPI space to transform the entire white and gray-matter surfaces into the distorted EPI space, and use these to define a surface-based searchlight.

High-field fMRI has been used to directly demonstrate columnar arrangements for

ocular dominance (Yacoub et al., 2007), orientation (Yacoub et al., 2008) and axis-of-motion (Zimmermann et al., 2011) in human visual cortex. Sub-millimetre resolution is required for direct demonstrations such as these, and the use of spin-echo (rather than gradient-echo) methods to reduce the PSF of the BOLD response and reduce the contribution from large, non-selective draining veins. The resolution used in the experiments in this thesis, combined with our use of gradient-echo imaging, precluded our ability to isolate individual cortical columns. However, our use of a higher resolution than that available in most 'standard' fMRI studies (1.5 mm isotropic compared with 3 mm isotropic) allowed us to increase the number of voxels available for classification in areas that may have had too few voxels in previous studies (e.g. MT+) (Kamitani and Tong, 2006), and also assess the impact of voxel size on classification (Gardner et al., 2006; Kriegeskorte et al., 2010).

Future Work

We have demonstrated the ability of multivariate classification methods for classifying direction of motion. Although we have also shown that classification for certain motion stimuli may be based on a coarse-scale signal rather than information arising from the distribution of columns or neurons, classification still has the potential to tells us about the processing of motion stimuli. Generalisation between different classes of stimuli can demonstrate a shared neural substrate, so investigating which stimuli will generalise to others, and which areas generate patterns that allow this, would be informative. For example, does the generalisation seen in MT+ for grating and plaid stimuli extend to plaids with different components, or between stimuli with very different spatial frequency profiles? For example, is generalisation possible between drifting grating stimuli and random dot stimuli with the same overall direction of drift? Serences and Boynton (2007b) were able to classify the perceived direction of an ambiguous motion stimulus from the patterns of motion arising from unambiguous motion stimuli in MT+, and it would be illuminating to see if this generalisation between actual-motion and other ambiguous or implied motion stimuli persists in other domains.

The results using plaid and paired-dot stimuli indicate how patterns of activity in different visual areas can rely on the perception of that stimulus by the observer. Future work could examine the classification of stimuli on a trial-by-trial basis, to see how changes in the percepts of stimuli (e.g. transparent versus coherent pattern motion) are reflected in changes in the behaviour of classifiers trained on patterns of activity from different visual areas. This also allows the decision itself about the stimulus to be clas-

sified, to find out where in the brain this decision itself is encoded (Hebart et al., 2012; Li et al., 2009; Serences and Boynton, 2007b).

Recent results have indicated that the classification of stimuli can be changed by experience, either improving in certain areas (Shibata et al., 2012) or altering the perceptual boundary between stimuli (Li et al., 2009). Future work could explore these changes in more detail for training using different stimuli and tasks to see whether different areas reflect concurrent changes. Changes in neural behaviour can also occur on a much shorter time-scale, namely after adaptation. Examining changes in classifier performance after adaptation could potentially reveal which areas are involved in these changes, although care may have to be taken when considering the difficulties in localizing fMRI-adaptation effects (Bartels et al., 2008).

Conclusions

Results such as generalisation between different stimulus types demonstrate the usefulness of classification methods, even if the scale of the signal being used by the classifier is currently still unknown. Successful generalization between different stimulus types indicates a sufficiently similar neural representation for those kinds of stimuli (Hong et al., 2011). The finding that classification can be affected by attentional effects (Kamitani and Tong, 2005, 2006) indicates that attention affects the neural representations of stimuli in a given visual area. Classification of stimuli can also be affected by training (Li et al., 2009; Shibata et al., 2012), which can be used to infer the neural changes involved with perceptual learning. These results do not depend on knowing the exact source of the signal being used in classification to make inferences about the neural representations involved. However, if definitive evidence about the fine-scale arrangement of feature-selective neurons is required, high-resolution imaging provides a more powerful approach (Bartels et al., 2008).

Developments in scanner technology are likely to lead to increasing improvements in both spatial (Yacoub et al., 2008) and temporal resolution (Feinberg et al., 2010). Increases in spatial resolution have already allowed the direct imaging of cortical columns in humans (Yacoub et al., 2007, 2008), and further developments will allow this trend to continue. Although the temporal resolution of fMRI is inherently limited by the hemodynamic response, increases in temporal resolution (via a reduction in TR) (Feinberg et al., 2010) may allow the response to be more accurately mapped, particularly the transient reduction in BOLD signal that follows neural activity (the 'initial dip'), which may be more closely linked to neural activity than the subsequent positive sig-

nal (Yacoub et al., 2001a). As well as offering these benefits, these developments also allow the possibility of a better understanding of the links between neural activity and hemodyanamics that lead to the BOLD response. This in turn should allow a better understanding of fMRI results (Kriegeskorte et al., 2010).

Multivariate classification methods have become popular in recent years as they offer the possibility of studying at very fine spatial (and more recently, temporal) scales, even using standard fMRI techniques (Tong and Pratte, 2012). However, the information being used by these methods is still unclear (Kriegeskorte et al., 2010): some experiments offer varying evidence for contributions for very fine scale information and information at much coarser scales (Freeman et al., 2011; Op de Beeck, 2010; Swisher et al., 2010), others indicate a role for the vasculature in transposing information into different spatial frequency bands of the signal (Gardner, 2010; Kriegeskorte et al., 2010). It may be the case that different experimental techniques emphasise information from different spatial scales (Clifford et al., published on-line 26 April 2011), or classification in individual subjects may be driven by information at different spatial scales (Misaki et al., 2012). Nevertheless, techniques such as these allow novel questions to be asked about neural representations in the brain.

References

DL Adams, LC Sincich, and JC Horton. Complete pattern of ocular dominance columns in human primary visual cortex. *The Journal of Neuroscience*, 27(39):10391–10403, 2007.

EH Adelson and JA Movshon. Phenomenal coherence of moving visual patterns. *Nature*, 300(5892):523–525, 1982.

TD Albright, R Desimone, and CG Gross. Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology*, 51(1):16–31, 1984.

K Amano, BA Wandell, and SO Dumoulin. Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *Journal of Neurophysiology*, 102(5):2704–2718, 2009.

D Apthorp, DS Schwarzkopf, C Kaul, B Bahrami, D Alais, and G Rees. Direct evidence for encoding of motion streaks in human visual cortex. *Proceedings of the Royal Society B: Biological Sciences*, 280(1752), 2013.

H Ban, TJ Preston, A Meeson, and AE Welchman. The integration of motion and disparity cues to depth in dorsal visual cortex. *Nature Neuroscience*, 15(4):636–643, 2012.

A Bartels, NK Logothetis, and K Moutoussis. fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. *Trends in Neurosciences*, 31(9):444–453, 2008.

E Bartfeld and A Grinvald. Relationships between orientation-preference pinwheels, cytochrome oxidase blobs, and ocular-dominance columns in primate striate cortex. *Proceedings of the National Academy of Sciences*, 89(24):11905–11909, 1992.

HGT Becker, M Erb, and T Haarmeier. Differential dependency on motion coherence in subregions of the human MT+ complex. *European Journal of Neuroscience*, 28(8): 1674–1685, 2008.

C Blakemore and FW Campbell. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology*, 203(1):237–260, 1969.

T Bonhoeffer and A Grinvald. Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature*, 353(6343):429–431, 1991.

RT Born and DC Bradley. Structure and function of visual area MT. *Annual Review of Neuroscience*, 28:157–189, 2005.

GM Boynton. Imaging orientation selectivity: decoding conscious perception in V1. *Nature Neuroscience*, 8(5):541–542, 2005.

GM Boynton and EM Finney. Orientation-specific adaptation in human visual cortex. *Journal of Neuroscience*, 23(25):8781–8787, 2003.

GM Boynton, SA Engel, GH Glover, and DJ Heeger. Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience*, 16(13): 4207–4421, 1996.

OJ Braddick and N Qian. The organization of global motion and transparency. In JM Zanker and J Zeil, editors, *Motion Vision - Computational, Neural, and Ecological Constraints*, chapter The Organization of Global Motion and Transparency, pages 85–112. Springer Verlag, Berlin Heidelberg New York, 2001.

OJ Braddick, JMD O'Brien, J Wattam-Bell, J Atkinson, T Hartley, and R Turner. Brain areas sensitive to coherent visual motion. *Perception*, 30(1):61–72, 2001.

KH Britten, MN Shadlen, WT Newsome, and JA Movshon. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *The Journal of Neuroscience*, 12(12):4745–4765, 1992.

KH Britten, MN Shadlen, WT Newsome, and JA Movshon. Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience*, 10:1157–1169, 1993.

GJ Brouwer and DJ Heeger. Decoding and reconstructing color from responses in human visual cortex. *Journal of Neuroscience*, 29(44):13992–14003, 2009.

GJ Brouwer and R van Ee. Visual cortex allows prediction of perceptual states during ambiguous structure-from-motion. *The Journal of Neuroscience*, 27(5):1015–1023, 2007.

A Burkhalter and DC Van Essen. Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *The Journal of Neuroscience*, 6(8):2327–2351, 1986.

RB Buxton, EC Wong, and LR Frank. Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magnetic Resonance in Medicine*, 39(6): 855–864, 2005.

EM Callaway. Structure and function of parallel pathways in the primate early visual system. *The Journal of Physiology*, 566(1):13–19, 2005.

FW Campbell, GF Cooper, and C Enroth-Cugell. The spatial selectivity of the visual cells of the cat. *The Journal of Physiology*, 203(1):223–235, 1969.

M Castelo-Branco, E Formisano, W Backes, F Zanella, S Neuenschwander, W Singer, and R Goebel. Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proceedings of the National Academy of Sciences*, 99(21): 13914–13919, 2002.

S Celebrini and WT Newsome. Neuronal and psychophysical sensitivity to motion signals in extrastriate area mst of the macaque monkey. *The Journal of Neuroscience*, 14(7):4109–4124, 1994.

D Chaimow, E Yacoub, K Ugurbil, and A Shmuel. Modeling and analysis of mechanisms underlying fMRI-based decoding of information conveyed in cortical columns. *NeuroImage*, 56(2):627–642, 2011.

C-C Chang and C-J Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:1–27, 2011. Software available at `http://www.csie.ntu.edu.tw/ cjlin/libsvm`.

Y Chen, P Namburi, LT Elliott, J Heinzle, CS Soon, MWL Chee, and JD Haynes. Cortical surface-based searchlight decoding. *Neuroimage*, 56(2):582–592, 2010.

MM Churchland, NJ Priebe, and SG Lisberger. Comparison of the spatial limits on direction selectivity in visual areas MT and V1. *Journal of Neurophysiology*, 93(3): 1235–1245, 2005.

CWG Clifford. Perceptual adaptation: motion parallels orientation. *Trends in Cognitive Sciences*, 6(3):136–143, 2002.

CWG Clifford, DJ Mannion, and JS McDonald. Radial biases in the processing of motion and motion-defined contours by human visual cortex. *Journal of Neurophysiology*, 102(5):2974–2981, 2009.

CWG Clifford, DJ Mannion, K Seymour, JS McDonald, and A Bartels. Are coarse-scale orientation maps really necessary for orientation decoding? *Journal of Neuroscience*, eLetter to the Editor, published on-line 26 April 2011.

D Cox and R Savoy. fMRI brain reading: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage*, 19(2):261–270, 2003.

W Curran and OJ Braddick. Speed and direction of locally-paired dot patterns. *Vision Research*, 40(16):2115–2124, 2000.

AM Dale, B Fischl, and MI Sereno. Cortical Surface-Based analysis* 1:: I. segmentation and surface reconstruction. *Neuroimage*, 9(2):179–194, 1999.

RL De Valois, DG Albrecht, and LG Thorell. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22(5):545–559, 1982a.

RL De Valois, EW Yund, and N Hepler. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, 22(5):531–544, 1982b.

GC DeAngelis and WT Newsome. Organization of disparity-selective neurons in macaque area MT. *The Journal of Neuroscience*, 19(4):1398–1415, 1999.

A Derrington and M Suero. Motion of complex patterns is computed from the perceived motions of their components. *Vision Research*, 31(1):139–149, 1991.

R Desimone and SJ Schein. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*, 57(3):835–868, 1987.

R Desimone and LG Ungerleider. Multiple visual areas in the caudal superior temporal sulcus of the macaque. *The Journal of Comparative Neurology*, 248(2):164–189, 1986.

ACM Diogo, JGM Soares, A Koulakov, TD Albright, and R Gattass. Electrophysiological imaging of functional architecture in the cortical middle temporal visual area of cebus apella monkey. *The Journal of Neuroscience*, 23(9):3881–3898, 2003.

RF Dougherty, VM Koch, AA Brewer, B Fischer, J Modersitzki, and BA Wandell. Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, 3(10):586–598, 2003.

R Dubner and SM Zeki. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research*, 35(2):528–532, 1971.

RO Duda, PE Hart, and DG Stork. *Pattern classification*. Pattern Classification and Scene Analysis: Pattern Classification. Wiley, 2001.

SO Dumoulin and BA Wandell. Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2):647–660, January 2008.

B Efron and RJ Tibshirani. *An introduction to the bootstrap.* Chapman & Hall, 1993.

SA Engel, GH Glover, and BA Wandell. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, 7(2):181–192, 1997.

F Fang, SO Murray, D Kersten, and S He. Orientation-tuned fMRI adaptation in human visual cortex. *Journal of Neurophysiology*, 94(6):4188–4195, 2005.

DA Feinberg, S Moeller, SM Smith, E Auerbach, S Ramanna, MF Glasser, KL Miller, K Ugurbil, and E Yacoub. Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS One*, 5(12):e15710, 2010.

B Fischl, MI Sereno, and AM Dale. Cortical Surface-Based analysis* 1:: II: inflation, flattening, and a Surface-Based coordinate system. *Neuroimage*, 9(2):195–207, 1999.

J Frahm, KD Merboldt, W Hänicke, A Kleinschmidt, and H Boecker. Brain or vein- oxygenation or flow? on signal physiology in functional mri of human brain activation. *NMR in Biomedicine*, 7(1-2):45–53, 1994.

J Freeman, GJ Brouwer, DJ Heeger, and EP Merriam. Orientation decoding depends on maps, not columns. *The Journal of Neuroscience*, 31(13):4792–4804, 2011.

CS Furmanski and SA Engel. An oblique effect in human primary visual cortex. *Nature Neuroscience*, 3(6):535–536, 2000.

JO Garcia and ED Grossman. Motion opponency and transparency in the human middle temporal area. *European Journal of Neuroscience*, 30(6):1172–1182, 2009.

JL Gardner. Is cortical vasculature functionally organized? *Neuroimage*, 49(3):1953–1956, 2010.

JL Gardner, P Sun, RA Waggoner, K Ueno, K Tanaka, and K Cheng. Contrast adaptation and representation in human early visual cortex. *Neuron*, 47(4):607–620, 2005.

JL Gardner, P Sun, K Tanaka, DJ Heeger, and K Cheng. Classification analysis with high spatial resolution fMRI reveals large draining veins with orientation specific responses. In *Society for Neuroscience Abstracts*, pages 640–614, 2006.

AN Garroway, PK Grannell, and P Mansfield. Image formation in NMR by a selective irradiative process. *Journal of Physics C Solid State Physics*, 7:L457–L462, 1974.

KR Gegenfurtner, DC Kiper, and JB Levitt. Functional properties of neurons in macaque area V3. *Journal of Neurophysiology*, 77(4):1906–1923, 1997.

Wilson S Geisler. Motion streaks provide a spatial code for motion direction. *Nature*, 400(6739):65–69, 1999.

MA Goodale and AD Milner. Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1):20–25, 1992.

K Grill-Spector and R Malach. fmr-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta psychologica*, 107(1):293–321, 2001.

K Grill-Spector, R Henson, and A Martin. Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1):14–23, 2006.

A Grinvald, E Lieke, RD Frostig, CD Gilbert, and TN Wiesel. Functional architecture of cortex revealed by optical imaging of intrinsic signals. *Nature*, 324:361–364, 1986.

C.G. Gross. Contribution of striate cortex and the superior colliculus to visual function in area MT, the superior temporal polysensory area and inferior temporal cortex. *Neuropsychologia*, 29(6):497–515, 1991.

K Guo, PJ Benson, and C Blakemore. Pattern motion is present in V1 of awake but not anaesthetized monkeys. *European Journal of Neuroscience*, 19(4):1055–1066, 2004.

SA Harrison and F Tong. Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238):632–635, 2009.

JD Haynes and G Rees. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8(5):686–691, 2005a.

JD Haynes and G Rees. Predicting the stream of consciousness from activity in human visual cortex. *Current Biology*, 15(14):1301–1307, 2005b.

JD Haynes and G Rees. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7):523–534, 2006.

JD Haynes, R Deichmann, and G Rees. Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature*, 438(7067):496–499, 2005.

S He, ER Cohen, and X Hu. Close correlation between activity in brain area MT/V5 and the perception of a visual motion aftereffect. *Current Biology*, 8(22):1215–1218, 1998.

MN Hebart, TH Donner, and JD Haynes. Human visual and parietal cortex encode visual choices independent of motor plans. *NeuroImage*, 63:1393–1403, 2012.

DJ Heeger and D Ress. What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience*, 3(2):142–151, 2002.

DJ Heeger, GM Boynton, JB Demb, E Seidemann, and WT Newsome. Motion opponency in visual cortex. *Journal of Neuroscience*, 19(16):7162–7174, 1999.

DJ Heeger, AC Huk, WS Geisler, and DG Albrecht. Spikes versus BOLD: what does neuroimaging tell us about neuronal activity? *Nature Neuroscience*, 3(7):631–632, 2000.

J. Hegdé and D.C. Van Essen. Selectivity for complex shapes in primate visual area V2. *Journal of Neuroscience*, 20(5):61–61, 2000.

K Hol and S Treue. Different populations of neurons contribute to the detection and discrimination of visual motion. *Vision Research*, 41(6):685–689, 2001.

SW Hong, F Tong, and AE Seiffert. Direction-selective patterns of activity in human visual cortex suggest common neural substrates for different types of motion. *Neuropsychologia*, 50(4):514–521, 2011.

DH Hubel and TN Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591, 1959.

DH Hubel and TN Wiesel. Integrative action in the cat's lateral geniculate body. *The Journal of Physiology*, 155(2):385–398, 1961.

DH Hubel and TN Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962.

DH Hubel and TN Wiesel. Shape and arrangement of columns in cat's striate cortex. *The Journal of Physiology*, 165(3):559–568, 1963.

DH Hubel and TN Wiesel. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28(2):229–289, 1965.

DH Hubel and TN Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1):215–243, 1968.

DH Hubel and TN Wiesel. Anatomical demonstration of columns in the monkey striate cortex. *Nature*, 221(182):747–50, 1969.

DH Hubel and TN Wiesel. Sequence regularity and geometry of orientation columns in the monkey striate cortex. *The Journal of Comparative Neurology*, 158(3):267–293, 1974.

DH Hubel and TN Wiesel. Ferrier lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 198: 1–59, 1977.

AC Huk and DJ Heeger. Pattern-motion responses in human visual cortex. *Nature Neuroscience*, 5(1):72–75, 2002.

AC Huk, D Ress, and DJ Heeger. Neuronal basis of the motion aftereffect reconsidered. *Neuron*, 32(1):161–172, 2001.

AC Huk, RF Dougherty, and DJ Heeger. Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, 22(16):7195–7205, 2002.

Y Kamitani and Y Sawahata. Spatial smoothing hurts localization but not information: pitfalls for brain mappers. *Neuroimage*, 49(3):1949–1952, 2010.

Y Kamitani and F Tong. Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8(5):679–685, 2005.

Y Kamitani and F Tong. Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology*, 16(11):1096–1102, 2006.

A Klein, J Andersson, BA Ardekani, J Ashburner, B Avants, MC Chiang, GE Christensen, DL Collins, J Gee, and P Hellier. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage*, 46(3):786–802, 2009.

A Kohn and JA Movshon. Neuronal adaptation to visual motion in area MT of the macaque. *Neuron*, 39(4):681–692, 2003.

A Kohn and JA Movshon. Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience*, 7(7):764–772, 2004.

H Kolster, R Peeters, and GA Orban. The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *The Journal of Neuroscience*, 30(29): 9801–9820, 2010.

N Kriegeskorte, R Goebel, and P Bandettini. Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (10):3863–3868, 2006.

N Kriegeskorte, R Cusack, and P Bandettini. How does an fMRI voxel sample the neuronal activity pattern: Compact-kernel or complex spatiotemporal filter? *Neuroimage*, 49(3):1965–1976, 2010.

SW Kuffler. Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16(1):37–68, 1953.

J Larsson and DJ Heeger. Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience*, 26(51):13128–13142, 2006.

J Larsson, MS Landy, and DJ Heeger. Orientation-selective adaptation to first-and second-order patterns in human visual cortex. *Journal of Neurophysiology*, 95(2):862, 2006.

HA Lee and SH Lee. Hierarchy of direction-tuned motion adaptation in human visual cortex. *Journal of Neurophysiology*, 107(8):2163–2184, 2012.

JY Lettvin, HR Maturana, WS McCulloch, and WH Pitts. What the frog's eye tells the frog's brain. *Proceedings of the IRE*, 47(11):1940–1951, 1959.

AG Leventhal. Relationship between preferred orientation and receptive field position of neurons in cat striate cortex. *The Journal of Comparative Neurology*, 220(4):476–483, 1983.

JB Levitt, DC Kiper, and JA Movshon. Receptive fields and functional architecture of macaque V2. *Journal of Neurophysiology*, 71(6):2517–2542, 1994.

S Li, SD Mayhew, and Z Kourtzi. Learning shapes the representation of behavioral choice in the human brain. *Neuron*, 62(3):441–452, 2009.

MS Livingstone and DH Hubel. Anatomy and physiology of a color system in the primate visual cortex. *The Journal of Neuroscience*, 4(1):309–356, 1984.

MS Livingstone and DH Hubel. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, 240(4853):740–749, 1988.

NK Logothetis. What we can do and what we cannot do with fMRI. *Nature*, 453(7197): 869–878, 2008.

NK Logothetis, J Pauls, M Augath, T Trinath, and A Oeltermann. Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843):150–157, 2001.

HD Lu, G Chen, H Tanigawa, and AW Roe. A motion direction map in macaque V2. *Neuron*, 68(5):1002 – 1013, 2010.

DC Lyon and JD Connolly. The case for primate V3. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729):625–633, 2012.

D Malonek, RBH Tootell, and A Grinvald. Optical imaging reveals the functional architecture of neurons processing shape and motion in owl monkey area MT. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 258(1352):109, 1994.

DJ Mannion, JS McDonald, and CWG Clifford. Discrimination of the local orientation structure of spiral glass patterns early in human visual cortex. *Neuroimage*, 46(2): 511–515, 2009.

P Mansfield. Multi-planar image formation using nmr spin echoes. *Journal of Physics C Solid State Physics*, 10:L55–L58, 1977.

D Marr and S Ullman. Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 211(1183):151–180, 1981.

JH Maunsell and DC van Essen. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of Neuroscience*, 3(12):2563–2586, 1983.

JH Maunsell, TA Nealey, and DD DePriest. Magnocellular and parvocellular contributions to responses in the middle temporal visual area (mt) of the macaque monkey. *The Journal of Neuroscience*, 10(10):3323–3334, 1990.

DJ McKeefry, JDG Watson, RSJ Frackowiak, K Fong, and SM Zeki. The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion. *Neuroimage*, 5(1):1–12, 1997.

RS Menon, S Ogawa, X Hu, JP Strupp, P Anderson, and K Ugurbil. Bold based functional mri at 4 tesla includes a capillary bed contribution: Echo-planar imaging correlates with previous optical imaging using intrinsic signals. *Magnetic Resonance in Medicine*, 33(3):453–459, 1995.

WH Merigan. Basic visual capacities and shape discrimination after lesions of extrastriate area v4 in macaques. *Visual Neuroscience*, 13(1):51–60, 1996.

WH Merigan, TA Nealey, and JH Maunsell. Visual effects of lesions of cortical area V2 in macaques. *The Journal of Neuroscience*, 13(7):3180–3191, 1993.

E Merriam, D Heeger, and J Freeman. Large scale biases for spirals and orientation in human visual cortex. In *Society for Neuroscience Abstracts*, 2012.

A Mikami, WT Newsome, and RH Wurtz. Motion selectivity in macaque visual cortex. ii. spatiotemporal range of directional interactions in MT and V1. *Journal of Neurophysiology*, 55(6):1328–1339, 1986.

AD Milner and MA Goodale. Two visual systems re-viewed. *Neuropsychologia*, 46(3): 774–785, 2008.

M Misaki, WM Luh, and PA Bandettini. The effect of spatial smoothing on fMRI decoding of columnar-level organization with linear support vector machine. *Journal of Neuroscience Methods*, 212(2):355–361, 2012.

VB Mountcastle, PW Davies, and AL Berman. Response properties of neurons of cat's somatic sensory cortex to peripheral stimuli. *Journal of Neurophysiology*, 20(4):374–407, 1957.

JA Movshon and WT Newsome. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *The Journal of Neuroscience*, 16(23):7733–7741, 1996.

JA Movshon, EH Adelson, MS Gizzi, and WT Newsome. The analysis of moving visual patterns. In C. Chagas, R. Gattass, and C. Gross, editors, *Pattern Recognition Mechanisms*, volume 54, pages 117–151. Rome: Vatican Press, 1985.

JA Movshon, TD Albright, GR Stoner, NJ Majaj, and MA Smith. Cortical responses to visual motion in alert and anesthetized monkeys. *Nature Neuroscience*, 6(1):3–4, 2003.

L Muckli, W Singer, FE Zanella, and R Goebel. Integration of multiple motion vectors over space: an fMRI study of transparent motion perception. *Neuroimage*, 16(4):843–856, 2002.

O Nestares and DJ Heeger. Robust multiresolution alignment of MRI brain volumes. *Magnetic Resonance in Medicine*, 43(5):705–715, 2000.

FH Netter. *Atlas of human anatomy*. Saunders, 2010.

WT Newsome and EB Pare. A selective impairment of motion perception following lesions of the middle temporal visual area MT. *The Journal of Neuroscience*, 8(6):2201–2211, 1988.

K Obermayer and GG Blasdel. Geometry of orientation and ocular dominance columns in monkey striate cortex. *The Journal of Neuroscience*, 13(10):4114–4129, 1993.

S Ogawa, TM Lee, AR Kay, and DW Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24):9868–9872, 1990.

H Okamoto, S Kawakami, H Saito, E Hida, K Odajima, D Tamanoi, and H Ohno. MT neurons in the macaque exhibited two types of bimodal direction tuning as predicted by a model for visual motion detection. *Vision Research*, 39(20):3465–3479, 1999.

NN Oosterhof, T Wiestler, PE Downing, and J Diedrichsen. A comparison of volume-based and surface-based multi-voxel pattern analysis. *Neuroimage*, 56(2):593–600, 2010.

HP Op de Beeck. Against hyperacuity in brain reading: Spatial smoothing does not hurt multivariate fMRI analyses? *Neuroimage*, 49(3):1943–1948, 2010.

CC Pack and RT Born. Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, 409(6823):1040–1042, 2001.

CC Pack, VK Berezovskii, and RT Born. Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. *Nature*, 414(6866):905–908, 2001.

CC Pack, MS Livingstone, KR Duffy, and RT Born. End-stopping and the aperture problem:: Two-dimensional motion signals in macaque V1. *Neuron*, 39(4):671–680, 2003.

R Pienaar, B Fischl, V Caviness, N Makris, and PE Grant. A methodology for analyzing curvature in the developing brain from preterm to adult. *International journal of imaging systems and technology*, 18(1):42–68, 2008.

BA Poser and DG Norris. Investigating the benefits of multi-echo EPI for fMRI at 7T. *Neuroimage*, 45(4):1161–1172, 2009.

BA Poser, MJ Versluis, JM Hoogduin, and DG Norris. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine*, 55(6):1227–1235, 2006.

WA Press, AA Brewer, RF Dougherty, AR Wade, and BA Wandell. Visual areas and spatial summation in human visual cortex. *Vision Research*, 41(10-11):1321–1332, 2001.

TJ Preston, S Li, Z Kourtzi, and AE Welchman. Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. *The Journal of Neuroscience*, 28(44):11315–11327, 2008.

N Qian and RA Andersen. Transparent motion perception as detection of unbalanced motion signals. II. physiology. *Journal of Neuroscience*, 14(12):7367, 1994.

N Qian and RA Andersen. V1 responses to transparent and nontransparent motions. *Experimental Brain Research*, 103(1):41–50, 1995.

N Qian, RA Andersen, and EH Adelson. Transparent motion perception as detection of unbalanced motion signals. I. psychophysics. *Journal of Neuroscience*, 14(12):7357, 1994.

M Raemaekers, MJM Lankheet, S Moorman, Z Kourtzi, and RJA van Wezel. Directional anisotropy of motion responses in retinotopic cortex. *Human Brain Mapping*, 30(12): 3970–3980, 2009.

G Rees, K Friston, and C Koch. A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neuroscience*, 3(7):716–723, 2000.

AW Roe and DY Ts'o. Visual topography in primate V2: multiple representation across functional stripes. *The Journal of Neuroscience*, 15(5):3689–3715, 1995.

H Saito, M Yukie, K Tanaka, K Hikosaka, Y Fukada, and E Iwai. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *The Journal of Neuroscience*, 6(1):145–157, 1986.

CD Salzman, CM Murasugi, KH Britten, and WT Newsome. Microstimulation in visual area MT: effects on direction discrimination performance. *The Journal of Neuroscience*, 12(6):2331–2355, 1992.

P Sapountzis, D Schluppeck, R Bowtell, and JW Peirce. A comparison of fMRI adaptation and multivariate pattern classification analysis in visual cortex. *Neuroimage*, 49 (2):1632–1640, 2010.

Y Sasaki. Processing local signals into global patterns. *Current Opinion in Neurobiology*, 17(2):132–139, 2007.

Y Sasaki, R Rajimehr, BW Kim, LB Ekstrom, W Vanduffel, and RBH Tootell. The radial bias: a different slant on visual orientation sensitivity in human and nonhuman primates. *Neuron*, 51(5):661–670, 2006.

T Schenk and RD McIntosh. Do we have independent visual streams for perception and action? *Cognitive Neuroscience*, 1(1):52–62, 2010.

D Schluppeck and SA Engel. Oblique effect in human MT+ follows pattern rather than component motion. *Journal of Vision*, 3(9):282, 2003.

B. Schölkopf, C.J.C. Burges, and A.J. Smola, editors. *Advances in Kernel Methods: Support Vector Learning*. MIT Press, Cambridge, MA, 1999.

F Segonne, AM Dale, E Busa, M Glessner, D Salat, HK Hahn, and B Fischl. A hybrid approach to the skull stripping problem in mri. *Neuroimage*, 22(3):1060–1075, 2004.

JT Serences and GM Boynton. Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron*, 55(2):301–312, 2007a.

JT Serences and GM Boynton. The representation of behavioral choice for motion in human visual cortex. *The Journal of Neuroscience*, 27(47):12893–12899, 2007b.

JT Serences, S Saproo, M Scolari, T Ho, and LT Muftuler. Estimating the influence of attention on population codes in human visual cortex using voxel-based tuning functions. *NeuroImage*, 44(1):223–231, January 2009.

K Seymour, CWG. Clifford, NK Logothetis, and A Bartels. The coding of color, motion, and their conjunction in the human visual cortex. *Current Biology*, 19(3):177–183, 2009.

K Shibata, LH Chang, D Kim, JE Náñez, Y Kamitani, T Watanabe, and Y Sasaki. Decoding reveals plasticity in V3A as a result of motion perceptual learning. *PloS one*, 7(8):e44003, 2012.

A Shmuel and A Grinvald. Functional organization for direction of motion and its relationship to orientation maps in cat area 18. *The Journal of Neuroscience*, 16(21): 6945–6964, 1996.

A Shmuel, G Raddatz, D Chaimow, NK Logothetis, K Ugurbil, and E Yacoub. Multiresolution classification analysis of ocular dominance columns obtained at 7 tesla from human V1: mechanisms underlying decoding signals. In *Society for Neuroscience Annual Meeting*, 2007.

A Shmuel, D Chaimow, G Raddatz, K Ugurbil, and E Yacoub. Mechanisms underlying decoding at 7 T: Ocular dominance columns, broad structures, and macroscopic blood vessels in V1 convey information on the stimulated eye. *Neuroimage*, 49(3): 1957–1964, 2010.

EP Simoncelli and DJ Heeger. A model of neuronal responses in visual area MT. *Vision Research*, 38(5):743–761, 1998.

JG Sled, AP Zijdenbos, and AC Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *Medical Imaging, IEEE Transactions on*, 17 (1):87–97, 1998.

A Smith, P Kosillo, and AL Williams. The confounding effect of response amplitude on MVPA performance measures. *Neuroimage*, 56(2):525–530, 2010.

AT Smith, MW Greenlee, KD Singh, FM Kraemer, and J Hennig. The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *The Journal of Neuroscience*, 18(10):3816–3830, 1998.

AT Smith, MB Wall, AL Williams, and KD Singh. Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience*, 23(2):561–569, 2006.

MA Smith, NJ Majaj, and JA Movshon. Dynamics of motion signaling by neurons in macaque area MT. *Nature Neuroscience*, 8(2):220–228, 2005.

RJ Snowden, S Treue, and RA Andersen. The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Experimental Brain Research*, 88(2):389–400, 1992.

GR Stoner and TD Albright. The interpretation of visual motion: evidence for surface segmentation mechanisms. *Vision Research*, 36(9):1291–1310, 1996.

P Sun, JL Gardner, M Costagli, K Ueno, RA Waggoner, K Tanaka, and K Cheng. Demonstration of tuning to stimulus orientation in the human visual cortex: A high-resolution fMRI study with a novel continuous and periodic stimulation paradigm. *Cerebral Cortex*, 2012.

JD Swisher, JC Gatenby, JC Gore, BA Wolfe, CH Moon, SG Kim, and F Tong. Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. *The Journal of Neuroscience*, 30(1):325–330, 2010.

OM Thomas, BG Cumming, and AJ Parker. A specialization for relative disparity in V2. *Nature Neuroscience*, 5(5):472–478, 2002.

SP Thompson. Optical illusions of motion. *Brain*, 3(3):289–298, 1880.

CJ Tinsley, BS Webb, NE Barraclough, CJ Vincent, A Parker, and AM Derrington. The nature of V1 neural responses to 2D moving patterns depends on receptive-field structure in the marmoset monkey. *Journal of Neurophysiology*, 90(2):930–937, 2003.

F Tong and MS Pratte. Decoding patterns of human brain activity. *Annual Review of Psychology*, 63(1):483–509, 2012.

F Tong, SA Harrison, JA Dewey, and Y Kamitani. Relationship between BOLD amplitude and pattern classification of orientation-selective activity in the human visual cortex. *NeuroImage*, 63(3):1212–1222, 2012.

RBH Tootell and N Hadjikhani. Where is 'dorsal V4' in human visual cortex? Retinotopic, topographic and functional evidence. *Cerebral Cortex*, 11(4):298–311, 2001.

RBH Tootell, JD Mendola, NK Hadjikhani, PJ Ledden, AK Liu, JB Reppas, MI Sereno, and AM Dale. Functional analysis of V3a and related areas in human visual cortex. *The Journal of Neuroscience*, 17(18):7060–7078, 1997.

RBH Tootell, N Hadjikhani, EK Hall, S Marrett, W Vanduffel, JT Vaughan, and AM Dale. The retinotopy of visual spatial attention. *Neuron*, 21(6):1409–1422, 1998.

DY Ts'o, M Zarella, and G Burkitt. Whither the hypercolumn? *The Journal of Physiology*, 587(12):2791–2805, 2009.

LG Ungerleider and M Mishkin. Two cortical visual systems. In M. A. Goodale D. J. Ingle and R. J. W. MansÞeld, editors, *Analysis of Visual Behavior*, page 549Ð586. MIT Press, Cambridge, MA, USA, 1982.

DC Van Essen and SM Zeki. The topographic organization of rhesus monkey prestriate cortex. *The Journal of Physiology*, 277(1):193–226, 1978.

DC Van Essen, JHR Maunsell, and JL Bixby. The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *The Journal of Comparative Neurology*, 199(3):293–326, 1981.

RJA Van Wezel and KH Britten. Motion adaptation in area mt. *Journal of Neurophysiology*, 88(6):3469–3476, 2002.

P van Zijl, J Hua, and H Lu. The bold post-stimulus undershoot, one of the most debated issues in fMRI. *NeuroImage*, 62(2):1092–1102, 2012.

MY Villeneuve, R Kupers, A Gjedde, M Ptito, and C Casanova. Pattern-motion selectivity in the human pulvinar. *Neuroimage*, 28(2):474–480, 2005.

MY Villeneuve, B Thompson, RF Hess, and C Casanova. Pattern-motion selective responses in MT, MST and the pulvinar of humans. *European Journal of Neuroscience*, 36 (6):2849–2858, 2012.

BA Wandell, AA Brewer, and RF Dougherty. Visual field map clusters in human cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456):693–707, 2005.

JDG Watson, R Myers, RSJ Frackowiak, JV Hajnal, RP Woods, JC Mazziotta, S Shipp, and SM Zeki. Area V5 of the human brain: evidence from a combined study using

positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, 3 (2):79, 1993.

MA Webster and RL De Valois. Relationship between spatial-frequency and orientation tuning of striate-cortex cells. *Journal of the Optical Society of America A*, 2(7):1124–1132, 1985.

MA Webster, D Kaping, Y Mizokami, and P Duhamel. Adaptation to natural facial categories. *Nature*, 428(6982):557–561, 2004.

Y Weiss, EP Simoncelli, and EH Adelson. Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6):598–604, 2002.

M Welicky, WH Bosking, and D Fitzpatick. A systematic map of direction preference in primary visual cortex. *Nature*, 379(6567):22, 1996.

TN Wiesel and DH Hubel. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, 29(6):1115–1156, 1966.

E Yacoub, A Shmuel, J Pfeuffer, V De Moortele, G Adriany, K Ugurbil, and X Hu. Investigation of the initial dip in fMRI at 7 tesla. *NMR in Biomedicine*, 14(7-8):408–412, 2001a.

E Yacoub, A Shmuel, J Pfeuffer, PF Van De Moortele, G Adriany, P Andersen, JT Vaughan, H Merkle, K Ugurbil, and X Hu. Imaging brain function in humans at 7 tesla. *Magnetic resonance in medicine*, 45(4):588–594, 2001b.

E Yacoub, TQ Duong, PF Van De Moortele, M Lindquist, G Adriany, SG Kim, K Uğurbil, and X Hu. Spin-echo fMRI in humans using high spatial resolutions and high magnetic fields. *Magnetic resonance in medicine*, 49(4):655–664, 2003.

E Yacoub, A Shmuel, N Logothetis, and K Ugurbil. Robust detection of ocular dominance columns in humans using hahn spin echo BOLD functional MRI at 7 tesla. *Neuroimage*, 37(4):1161–1177, 2007.

E Yacoub, N Harel, and K Ugurbil. High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences*, 105(30):10607–10612, 2008.

O Yamashita, M Sato, T Yoshioka, F Tong, and Y Kamitani. Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *NeuroImage*, 42(4):1414 – 1429, 2008.

SM Zeki. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *The Journal of Physiology*, 236(3):549–573, 1974.

SM Zeki. The response properties of cells in the middle temporal area (area MT) of owl monkey visual cortex. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 207(1167):239–248, 1980.

SM Zeki. The distribution of wavelength and orientation selective cells in different areas of monkey visual cortex. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 217(1209):449–470, 1983.

SM Zeki. A century of cerebral achromatopsia. *Brain*, 113(6):1721–1777, 1990.

SM Zeki. Improbable areas in the visual brain. *Trends in Neurosciences*, 26(1):23–26, 2003.

J Zihl, D Von Cramon, and N Mai. Selective disturbance of movement vision after bilateral brain damage. *Brain*, 106(2):313–340, 1983.

J Zimmermann, R Goebel, F De Martino, PF van de Moortele, D Feinberg, G Adriany, D Chaimow, A Shmuel, K Uğurbil, and E Yacoub. Mapping the organization of axis of motion selective features in human area MT using high-field fMRI. *PLoS ONE*, 6 (12):e28716, 12 2011.